Automotive Imbalance Dataset Analysis and Solution using Deep Learning Algorithm

M S Sunitha Patel[1], Srinath S[2]

[1]*Assistant Professor, Dept. of CSE, ATME College of Engineering, Mysuru, Karnataka, India, mssunithapatel@gmail.com*
[2]*Assosciate Professor, Dept. of CSE, SJCE, Mysuru, Karnataka, India srinath@sjce.ac.in*

Abstract.

A very basic stage and a very significant part of achieving effective results for the specified challenge is to train the machine learning algorithms with data. The primary skill in data science engineering is gathering raw data and building datasets, pre-processing and annotating for the given challenge. In machine learning-based problem-solving strategies, creating the correct dataset for automotive image processing takes around half the time. The varied weather conditions, road conditions, driving conditions, and other factors make automotive image processing more difficult. The collection of imbalanced datasets during image dataset construction is a common difficulty in deep learning models for multiclass analysis. The purpose of this study is to demonstrate the process of creating, testing, training, validating, and evaluating a dataset for three distinct vehicle classes: bus, car, and truck. The Convolutional Neural Networks model is used to evaluate performance and represented through Receiver Operating Characteristics (ROC) for balanced and unbalanced datasets (CNN). The ability of image augmentation to increase training performance on imbalanced datasets has been proved experimentally.

Keywords. Imbalance dataset, image data science, deep learning, CNN, Image Augmentation.

1. INTRODUCTION

The quality of data sets is determined by the problem statement, the product being utilized, the development time, and the development cost. To produce a fully autonomous car, both the dataset quality and the number of training images necessary must be exceedingly high. . Thus, for the given problem, it is vital to understand the type of data that will be utilised for training in order to avoid incurring extra development costs, lead times, and calculation time. Numerous automotive image processing systems require a fast response time [1] in order to drive the vehicle within milliseconds. Active safety systems such as Adaptive Cruise Control (ACC) [2], Lane Keeping Assistance (LKA) [3], and Collision Assistance (CA) [4], among others, require a very fast response time to avert catastrophic incidents. Another important consideration is that due to the necessity to research and train the Region Of Interest, the datasets required to train ACC, LKA, and CA may differ (ROI). The principal use of image processing in this situation may be to estimate the distance between two vehicles because the fundamental function of ACC is to manage the vehicle's speed and the spacing between vehicles ahead. If an unintended lane change occurs in LKA, the vehicle control system should inform the driver. So, in LKA, image data must be gathered that suits unintentional lane change circumstances such as sudden lane mark changes, lane mark fading, low visibility of lane markers owing to fog conditions, and so on. Several automotive image processing applications, such as Parking Assistance (PA) and Night Vision Camera Assistance (NVCA) [5], require a response time of less than one second. Due to the fact that the output of NVCA and PA systems is used by drives to perform the next control action, meeting the response time for people should be achievable. Because human beings govern the next action, the image's clarity is critical. NVCA places a greater emphasis on images recorded at night to aid drivers in the event of a predicted risk. It is self-evident that automotive image processing has more complex requirements in order to satisfy the diverse needs of end users. Automotive firms have a lot of opportunity to create cost-effective and trustworthy solutions with deep learning. The paper is organised into four pieces. The first portion describes the rationale for the current work, followed by difficulties and solutions in automotive image datasets, as well as literature studies. The third piece covers the approach used to solve the problem, the fourth section explains the simulation results for different dataset sizes, and the last section explains the conclusion and possible future study.

2. BACKGROUND

The number of cameras in modern vehicles has expanded in order to improve driver safety and comfort. In sophisticated feature automobiles, the number of cameras used can be more than ten [6] [7], and modern high safety vehicles contain numerous safety and driver comfort features. As the number of cameras used grows, image processing and data analysis tools will face additional obstacles. Because cameras are mounted on a vehicle that is subjected to dynamic conditions such as vibration, dust, and rain [8,] the issues begin with picture acquisition. The acquired images will include a lot of noise and unnecessary data due to the dynamic situations. Following image acquisition, image enhancement [9] is used to improve image quality such as resolution, colour rendition, and so on, depending on the noise level in the collected images. After the image has been improved, image segmentation is a crucial stage in the dataset production process. The data science engineer must capture the right objects or features in the datasets to train the machine learning model. Image segmentation mainly deals with the required 'object detection and classification,' so the data science engineer must capture the right objects or features in the datasets to train the machine learning model. To have a safe driving car, for example, sophisticated vehicle driver assistance systems such as drive less vehicles must detect which area in the acquired image or video is safe to drive the vehicle by analysing environments [10]. The final step in image processing is'representation and description,' which is the required section of the image that can be used for further processing after object recognition and calcification. However, for automobile safety applications, a high level of representation and description is essential. Pixel is the lowest level of feature that provides required information in representation and description.

Vehicle type categorization is a classic example of automotive image processing that is required in several of the vehicle applications listed above. The first step in applying deep learning to classify vehicles is to gather the appropriate images and develop datasets. Datasets can be built from scratch or acquired from available sources. There are research papers available for vehicle classification using CNN [11],[12] in which the researchers created their own dataset; in a few cases [13],[14], the researchers created the dataset for the vehicle's rear part; and in another few cases [15], VeRi-776[16], the researchers used vehicle datasets available in open source PKU-VD [15], VeRi-776[16]. To gain access to open source datasets, as indicated previously, researchers must obtain permission from the owners. However, in order to obtain consent from the proprietors, certain time delays or even no response may be seen. Additionally, other vehicle image datasets are freely accessible to the public. The primary difficulty in creating automobile picture databases is obtaining the necessary number of photographs for various applications. For example, photos or videos of various sorts of road lanes, vehicle types, and environmental conditions (rain, snow, dusty, etc.) are required to train the deep learning algorithms for LKS. The major issues in the automotive image datasets are:

1. Availability of image datasets for all the required features to train the model.
2. Availability of required image quantity to train the model.
3. Availability of appropriate images for real-time vehicle driving scenarios
4. Availability of appropriate images for different environmental conditions

All of these factors combine to create a significant common challenge: handling imabalance datasets. Imbalance data set occurs in classification problems during performance analysis of multiple classes with the uneven quantity of datasets. To overcome issues seen during imbalanced datasets, image augmentation can be utilized. In this paper, multiclass vehicle classification like bus, car, and truck has been analyzed in three different scenarios of image datasets. The main reason to take the 3 mentioned classes is that getting the required quantities of images for buses and trucks in open access is quite difficult.

3. METHODOLOGY

CNN architecture has been used to analyze the impact on different image dataset sizes and their performance on the machine learning algorithm. The methodology of dataset modeling, training, and validation for CNN architecture has been explained as follows.

3.1 MODELING DATASETS

Once the challenges inherent in automotive image processing are comprehended, image data collection for the specified task must begin. The primary issue with automobile image collections is a lack of required images in open-source. As previously stated, there are a few open sources accessible. However, for the chosen topic, which involves three distinct types of vehicles, while there are several vehicle datasets accessible, photos for bus and truck datasets were not available in sufficient quantities. As mentioned in table-1, car image datasets are gathered with 5000 images where has for bus and trucks managed to gather 2900 and 3400 images respectively from open source. To increase the image dataset size, a data augmentation approach has been utilized by using existing images. There are various data augmentation techniques are available like padding, cropping, flipping in different directions, etc. Depending on the problem and available images one can create additional images for the data sets. By using the image augmentation technique additional images for bus and truck has been created for 2100 and 1600 images respectively. Two examples of image augmentation have been shown in Figure 3.1 and Figure 3.2.

Tabel-1: Dataset Scenarios for Analysis

Dataset Scenarios	Number of Bus images	Number of Car images	Number of Truck images
1.Minimal Images	2900	2900	2900
2.Imbalance Images	2900	5000	3400
3. Added Augmented Images to make 5000 equal images	2100	0	1600

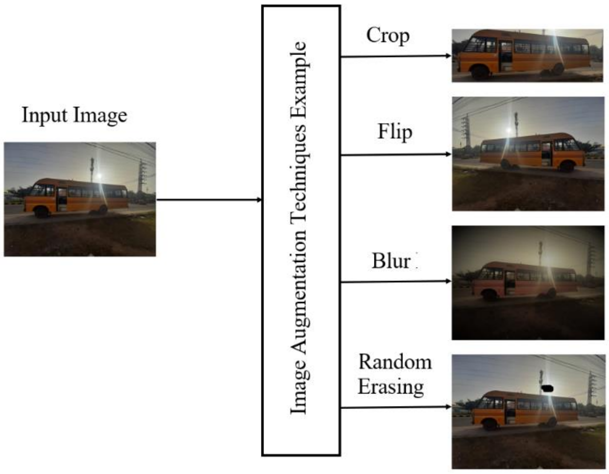

Figure 3.1 Image augmentation examples for bus image

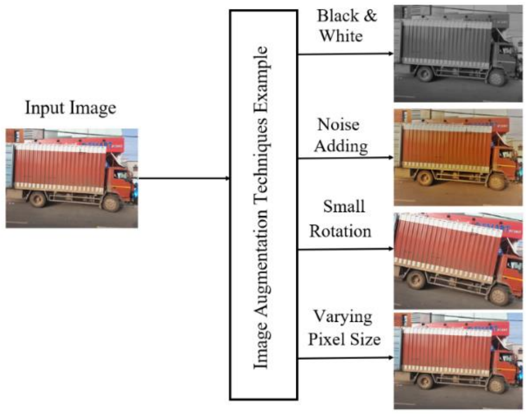

Figure 3.2 Image augmentation example for truck image

3.2 TRAINING, TESTING, AND VALIDATION DATASETS

As mentioned in Figure 3.3, total images in the datasets are segregated as training, testing, and validation. However, in the current scenario, only training and testing datasets in the ratio of 80% and 20% have been implemented. Sometimes validation datasets can also be used to fine-tune the learning algorithms.

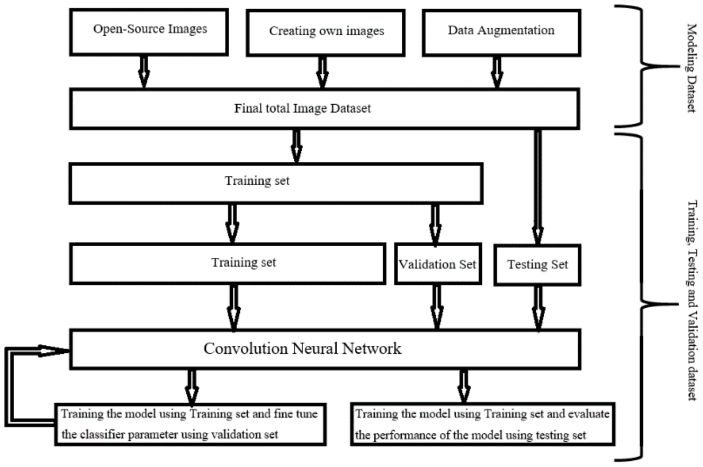

Figure 3.3 Testing and training for CNN

3.3 CNN ARCHITECTURE

A three-layer convolution neural network has been implemented as shown in Figure 3.4. An experiment has been carried out using Python 3.7.9, with IDE PyCharm and OpenCV Library. Three different datasets as mentioned in the table-1 have been created separately and three different inputs have been provided for the CNN model. Image size of 64 x 64 has been normalized and feature extraction including filtering, image segmentation, and image enhancement has been taken care of from the CNN model. CNN models consist of convolution layer, pooling layer, fully connected layer, and output layer. The first convolution layer consists of 32 filters and followed by 64 filters in the second and third layers. Padding has been considered to maintain input and output image size. Max pooling and Average pooling are widely used in the research article however max-pooling has been considered because of its benefits in capturing maximum value in a feature map. The output of max pooling with window size 2 and stride size 2 is fed to the size 2 of the fully connected layer. The fully connected layer is the deep layer that is used for classification. Finally, the fully connected layer passes its input to softmax with 2 layers for the final image classified output. In the current work, three different datasets are passed with fifty epochs for each scenario to train the learning algorithm. The output of each scenario is executed in three different phases and results have been captured and explained in the next section.

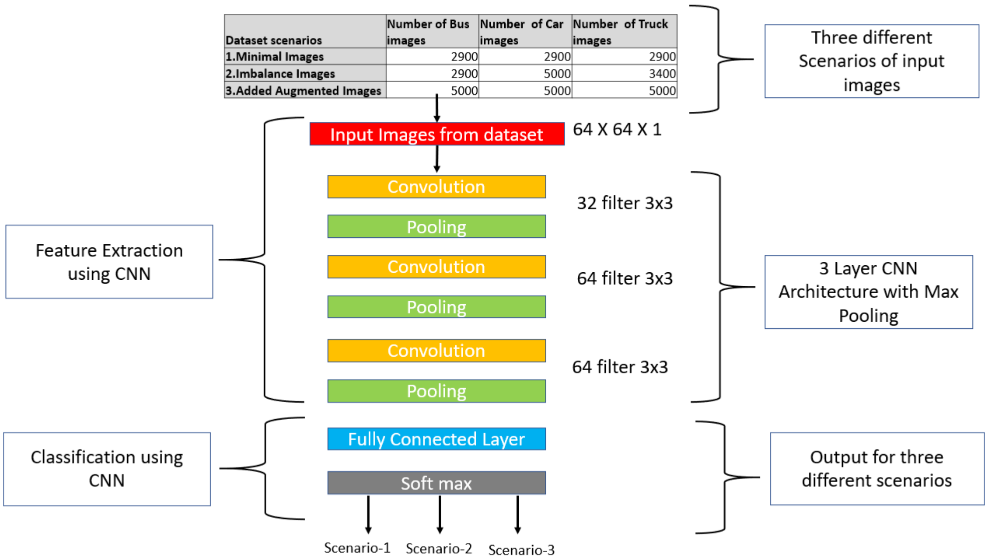

Figure 3.4 CNN Architecture

4. Result and Discussion

Performance evaluation of different dataset scenarios can be analyzed using accuracy as mentioned in equation 1. The result of the ROC curve is as shown in Figure 4.1, it is quite evident that an imbalanced dataset gives low performance compared to the balanced dataset. The main reason the performance of the imbalance dataset is deteriorated because of false negative and false positive will increase for bus and truck class. Also, it can be observed that after image augmentation the output ROC is increased from 88% to 98% scenario-1 to scenario-3.

$$Accuracy = \frac{(\text{True Positive} + \text{True Negative})}{(\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative})} \qquad (1)$$

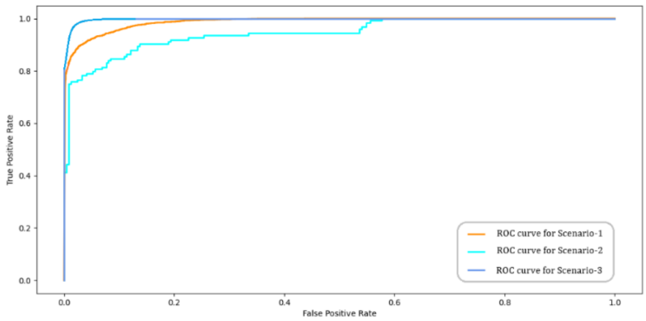

Figure 4.1 ROC curve for all three Scenarios

From Figure 4.2 as well it is quite evident that for the imbalance dataset error rate is more compared to the balanced dataset. After image augmentation, the error rate has reduced by nearly 10% compared to scenario-1 and scenario-3.

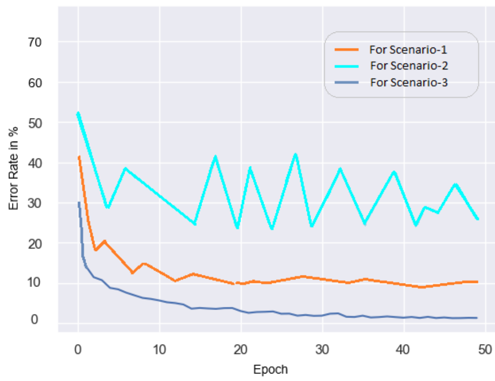

Figure 4.2 Performance graph for different scenario

5. FUTURE WORK

Adding images by utilizing the image augmentation technique has improved the performance of the CNN training model. The next part of our research work is to have a deep study on the different image augmentation techniques and their impact on performance. So future work will be concentrated on a detailed study of padding, cropping, flipping, rotation, etc, and its performance evaluation for different automotive functional conditions LKS, ACC, CA, etc.

6. REFERENCES

[1] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W.v. Seelen, “An Image Processing System for Driver Assistance,” in Proceedings of the IEEE Conference on Intelligent Vehicles, Stuttgart, Germany, 1999, pp. 481–486, IEEE

[2] M. M. Trivedi, T. Gandhi and J. McCall, "Looking-In and Looking-Out of a Vehicle: Computer-Vision-Based Enhanced Vehicle Safety," in IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 1, pp. 108-120, March 2007, DOI: 10.1109/TITS.2006.889442.

[3] Fei Han, Weiwen Deng, Sumin Zhang, Bei Ren and Ying Wang, AVision-based Forward Collision Warning System Developed under virtual Environment, SAE International Journal of TransportationSafety, Volume 4, Issue 2, July 2014, DOI: 10.4271/2014-01-0754.

[4] Shih-Shinh Huang, Chung-Jen Chen, Pei-Yung Hsiao and L. -C. Fu, "On-board vision system for lane recognition and front-vehicle detection to enhance driver's awareness," IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004, pp. 2456-2461 Vol.3, DOI: 10.1109/ROBOT.2004.1307429.

[5] J. McCall and M. M. Trivedi, “Human behavior based predictive brake assistance,” in Proc. IEEE Intelligent Vehicles Symphonium, Jun 2006.

[6] Rafael C. Gonzalez, Richard E Woods, Digital Image Processing, Third Edition, Dorling Kindersley (India) Pvt. Ltd,2014.

[7] M S Sunitha Patel and Dr.Srinath S, Soft Computing Approaches for Automotive Image Processing: Opportunities & Challenges, Advances in Automation, Signal Processing, Instrumentation, and Control, Scopus Indexed Springer Proceedings – Lecture Notes in Electrical Engineering 2020.

[8] Fei Han, Weiwen Deng, Sumin Zhang, Bei Ren and Ying Wang, AVision-based Forward Collision Warning System Developed under virtual Environment, SAE International Journal of TransportationSafety, Volume 2, Issue 2, July 2014

[9] Yunbo Rao, Leiting Chen, A Survey of Video Enhancement Techniques, Journal of Information Hiding and Multimedia Signal ProcessingVolume 3, Number 1, January 2012.

[10] J. McCall and M. M. Trivedi, Video based lane estimation and tracking for driver assistance: Survey, algorithms, and evaluation, IEEE Trans.Intell. Transp. Syst., vol. 7, no. 1, pp. 20–37, Mar. 2006.

[11] M. A. Hedeya, A. H. Eid, and R. F. Abdel-Kader, “A super learner ensemble of deep networks for vehicle-type classification,” IEEE Access, 2020.

[12] Abhishek N, S Gopalswamy, S Rathinam, “Vision based Techniques for Identifying Emergency Vehicles, SAE Technical Paper, 2019-01-0889, April 2019.

[13] Y. Lou, Y. Bai, J. Liu, S. Wang, and L. Duan, ``VERI-wild: A large dataset and a new method for vehicle re-identi_cation in the wild," inProc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 3235_3243.

[14] M. Naphade, S. Wang, D. C. Anastasiu, Z. Tang, M.-C. Chang, X. Yang, Y. Yao, L. Zheng, P. Chakraborty, C. E. Lopez, and. Sharma, ``The 5th AI city challenge," in Proc. IEEE/CVF Conf. Comput.Vis. Pattern Recognit. Workshops, Jun. 2021.

[15] Xinchen Liu, Wu Liu, Tao Mei, and Huadong Ma. A deep learning-based approach to progressive vehicle reidentification for urban surveillance. In Proc., ECCV, pages 869–884, 2016.

[16] Hongye Liu, Yonghong Tian, Yaowei Yang, Lu Pang, and Tiejun Huang. Deep relative distance learning: Tell the difference between similar vehicles. In Proc. CVPR, pages 2167–2175, 2016.

[17] https://www.kaggle.com/brsdincer/vehicle-detection-set

[18] https://data.mendeley.com/datasets/pwyyg8zmk5/2

[19] http://ai.stanford.edu/~jkrause/cars/car_dataset.html

[20] https://github.com/datacluster-labs/Indian-Vehicle-Image-Dataset

[21] https://www.gti.ssr.upm.es/data/Vehicle_database.html

[22] https://data.mendeley.com/datasets/pwyyg8zmk5/2

Biographies

M S Sunitha Patel received a Bachelor's Degree in computer engineering from Visvesvaraya Technological University, Belgaum in 2005, a Master's Degree from Visvesvaraya Technological University, Belgaum in 2010. She is currently working as an Assistant Professor at the Department of Computer Engineering, ATME College of Engineering, Mysuru, Karnataka. Her research areas include image processing, machine learning, data mining and deep learning.

Dr. Srinath.S is working as an Associate Professor, Department of Computer Science and Engineering, Sri Jayachamarajendra College of Engineering, Mysuru. He obtained his B.E Degree in Computer Science and Engineering from Siddaganga Institute of Technology in 1995. He later obtained his Masters in Software Engineering in the year 2002 and topped the university by securing first rank and gold medal from the esteemed Visvesvaraya Technological University, Belgaum. He obtained his Ph. D Degree in Computer Science for his thesis “Investigation and Development of Robust algorithms for recognizing and analyzing printed and hand embossed Kannada Braille characters: Novel approaches” from the University of Mysore in the year 2015. He has 25 years of rich experience in the field of teaching.