International Conference on Computing and Machine Learning (CML 2024)

Safe and Efficient Traffic Obstacle Detection using Fine-tuned Modified Faster R-CNN

(Paper ID 664)

Authors:

Jyoti Madake, Shruti Korpade, Hemal Kulkarni, Rahul Kumbhar, Shripad Bhatlawande

Vishwakarma Institute of Technology, Pune

❖Contents

- Introduction
- Literature Survey
- Research Gap
- Novelty
- Methodology
- Results
- Advantages and Limitations
- Conclusion

INTRODUCTION:

➔ Challenges in recognizing road obstacles

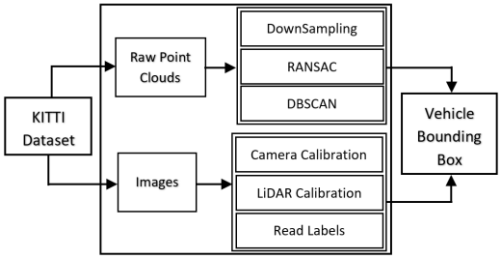
➔ Identification and Prediction of obstacles to avoid traffic jam

➔Traffic Obstacle Detection by Fine-tuning Faster R-CNN

LITERATURE SURVEY

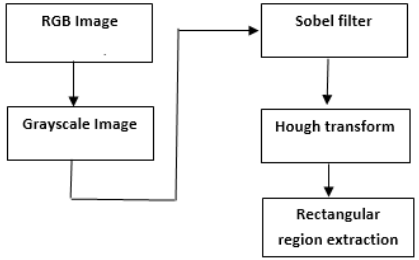
Visualization of 3D Point Clouds for Vehicle Detection Based on LiDAR and Camera Fusion
[1]

- Detection of vehicles using Camera and Lidar data.
- RANSAC, DBSCAN Algorithms used for visualization and clustering of the LiDAR point cloud.



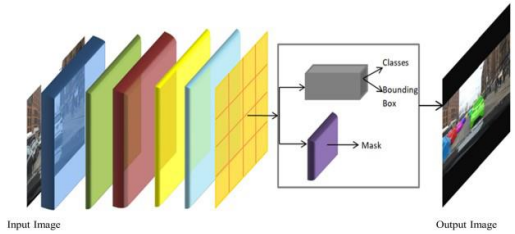
Preprocessed Faster RCNN for Vehicle Detection
[2]

- Uses Sobel edge operator and Hough Transform for lane detection.
- Extracts a Rectangular region from lane coordinates for reduced ROI.



Vehicle Detection through Instance Segmentation using Mask R-CNN for Intelligent Vehicle System
[3]

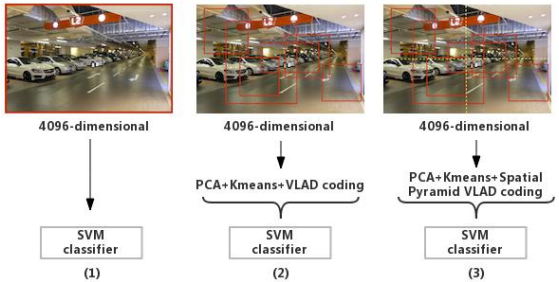
- Implementation of Mask R-CNN using transfer learning for vehicle detection.
- Segmentation-based approach adopted for precise vehicle identification.



Traffic scene recognition based on deep CNN and VLAD spatial pyramids

[4]

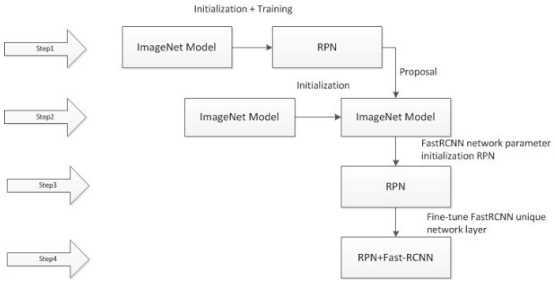
- Vector of Locally Aggregated Descriptors (VLAD) encoding is used to capture global contextual information.
- CNN features of image patches are encoded using VLAD Spatial pyramids.



Vehicle type detection based on deep learning in traffic scene

[5]

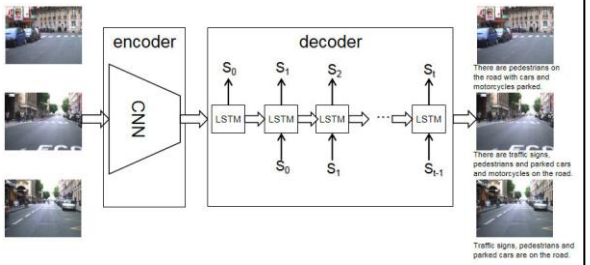
- Application of Faster RCNN framework and improvisation on RPN networks.
- Detection and recognization of three common types of vehicles in traffic scenes.



The traffic scene understanding and prediction based on image captioning

[6]

- The paper introduces an image captioning model.
- The solution solves feature fusion, general object recognition, and low-level semantic understanding problems.



❖Research Gap

- **Unable to detect road obstacles**
- **Challenges like Occlusion, Illumination Change, Response**
- **Expensive Real-time detection systems**
- **Low Accuracy**

❖Novelty

Approach: Fine tuning of Faster R-CNN

Real-time Detection and decision-making

Lowest Processing Time and Training Loss

Highly Accurate Lightweight System

METHODOLOGY

Dataset and Preprocessing:

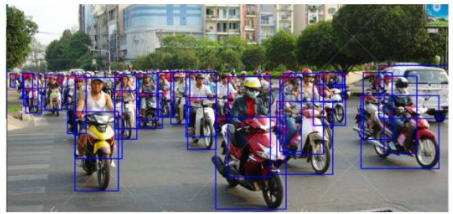



Sr. No.	Classes	No. of Training Images	No. of Testing Images
1	Pedestrians	944	235
2	Traffic Signs	551	162
3	Cars	1015	177
4	Bikes	331	82
5	Zebra Crossing	176	36
	Total	**3017**	**692**

- ➔ Dataset consists of 3709 RGB images.
- ➔ Algorithm uses data transformations for feature extraction using a ResNet-50 model.

System Block Diagram

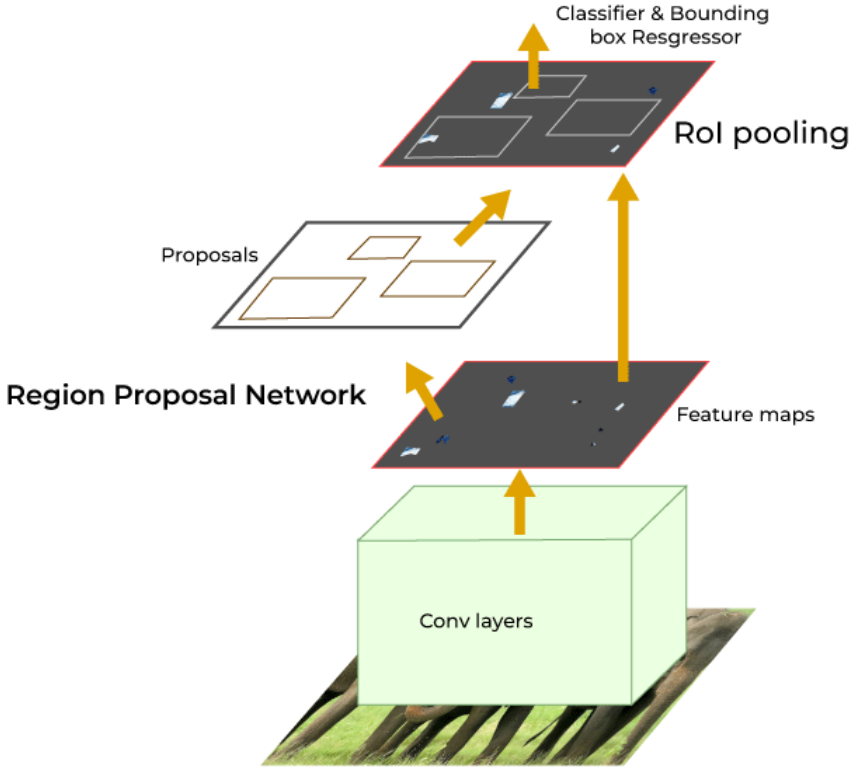
Input: Video Stream or Image → Train ResNet-50 model on custom dataset, for classification → Evaluate performance of model, to correctly classify images from test dataset → Demonstrate use of the model to predict class of object in image → Use of Faster R-CNN to draw bounding boxes around objects in video → Output: Traffic Object Detection Results (Bounded Boxes)



Faster R-CNN Architecture

Faster R-CNN architecture consists of two components:

1) **Region Proposal Network (RPN)**
2) **Fast R-CNN detector**

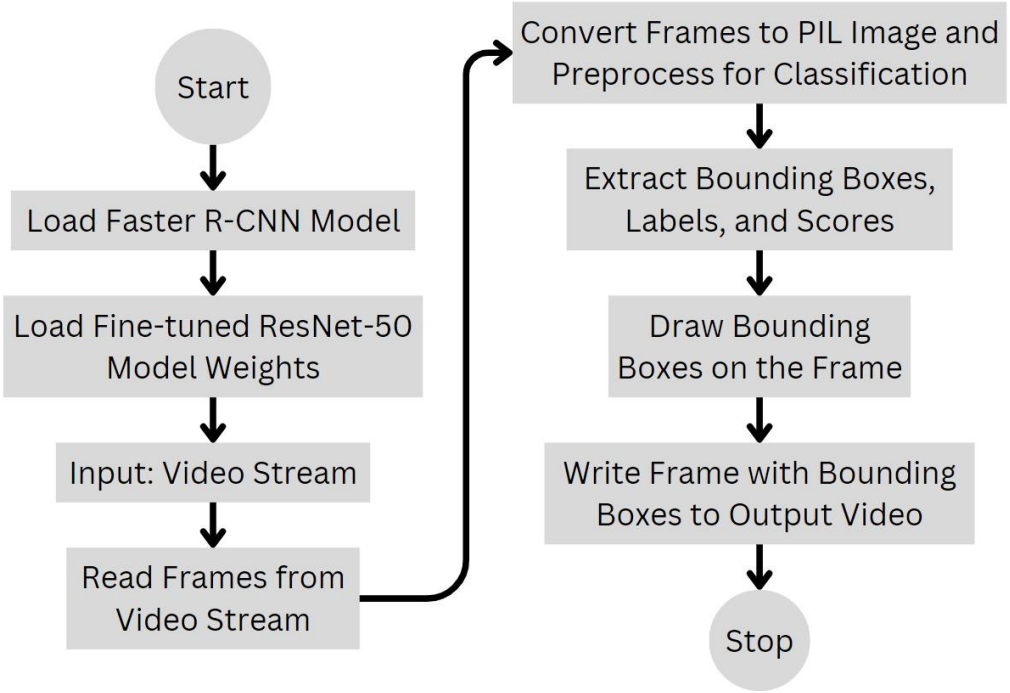


Credits for Image:
https://www.geeksforgeeks.org/faster-r-cnn-ml/

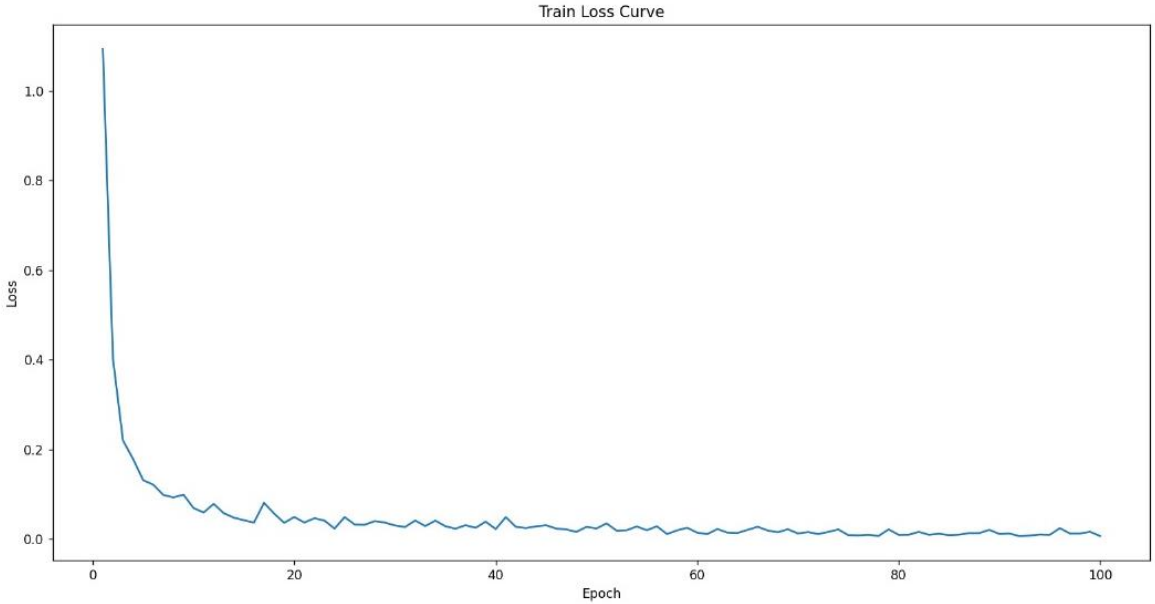
Proposed Model :

★ The algorithms are as follows:

- ResNet50
- Faster RCNN
- Stochastic Gradient Descent
- Cross Entropy
- Fine Tuning



Results and Discussion:

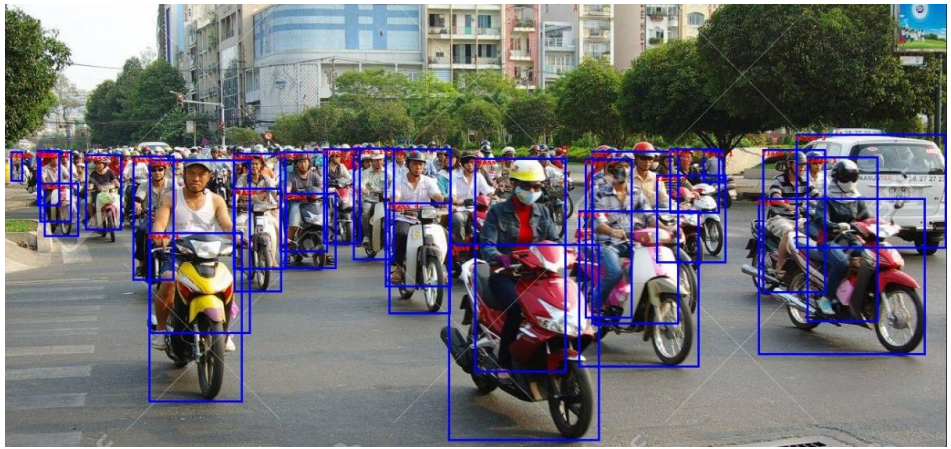


No. of Epochs	Test Accuracy	Final Training Loss	Process Inference: 156 milliseconds
100	98.2659%	0.0326	'ASUS TUF dash Gaming F15 Core i5 12th Gen 12500H - (8 GB/512 GB SSD/4 GB Graphics/NVIDIA GeForce RTX 3050)' machine

Results and Discussion:

INPUT:

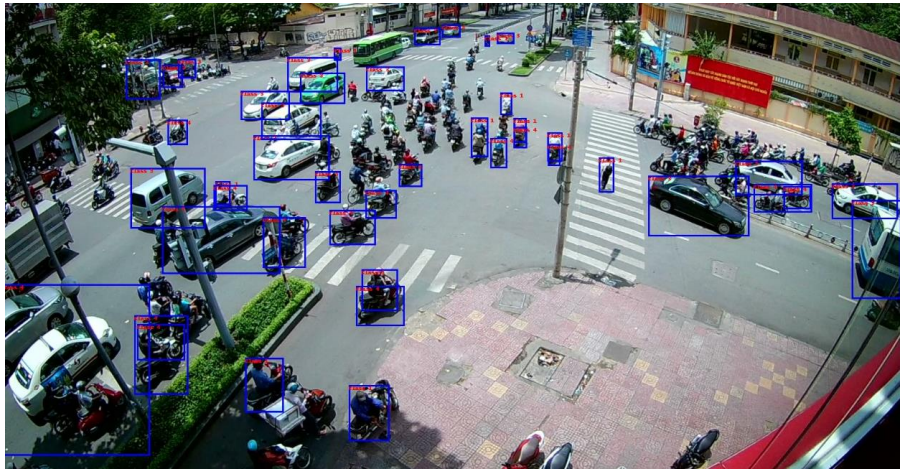
OUTPUT:



Results and Discussion:

SCENARIO 1:

SCENARIO 2:



Advantages

- ➔ Precise localization of bikes, cars, and pedestrians.
- ➔ High Accuracy: 98.27%
- ➔ Inference Time of 0.156 milliseconds
- ➔ Real-time decision-making

Limitations

- ➔ Blurred or noisy input images can leads to wrong prediction.
- ➔ Semantic Segmentation, Lane Detection, Traffic Scene Graph Understanding
- ➔ Increasing no. of classes in the dataset for detection of other obstacles.

Comparison of Proposed System:

Model	Accuracy (mAP)	Model Size (MB)	Dataset	Bounding box Loss
Proposed Model	**91%**	**92**	**Customized**	**0.0326**
Faster R-CNN (Resnet101)	85.60%	200	COCO, Pascal VOC2012	0.12
Fast R-CNN	70%	150	Pascal VOC2007	Smooth L1
R-CNN	66%	75	Pascal VOC2007	Smooth L1
YOLOv2 (Darknet19)	78.60%	35	Pascal VOC2007	0.057
SSD512 (VGG16)	80%	25	COCO, Pascal VOC2012	Smooth L1
RetinaNet	73.80%	75	COCO	Focal Loss

Conclusion

- ➔ Fine-tuned Modified Faster R-CNN system detects traffic obstacles with **98.27% accuracy**.
- ➔ Proposed Model is a lightweight system having **92 MB model size**.
- ➔ Final **Bounding Box training loss** was calculated as **0.0326**.
- ➔ The proposed system has an **Inference time of 156 milliseconds**.

❖ References:

1. J. Madake, R. Rane, R. Rathod, A. Sayyed, S. Bhatlawande and S. Shilaskar, "Visualization of 3D Point Clouds for Vehicle Detection Based on LiDAR and Camera Fusion," 2022 OITS International Conference on Information Technology (OCIT), Bhubaneswar, India, 2022, pp. 594-598. [1]

2. Manana, Mduduzi, Chunling Tu, and Pius Adewale Owolawi. "Preprocessed faster RCNN for vehicle detection." In 2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC), pp. 1-4. IEEE, 2018. [2]

3. Ojha, Apoorva, Satya Prakash Sahu, and Deepak Kumar Dewangan. "Vehicle detection through instance segmentation using mask R-CNN for intelligent vehicle system." In 2021 5th international conference on intelligent computing and control systems (ICICCS), pp. 954-959. IEEE, 2021. [3]

4. Wu, Fang-Yu, Shi-Yang Yan, Jeremy S. Smith, and Bai-Ling Zhang. "Traffic scene recognition based on deep CNN and VLAD spatial pyramids." In 2017 International Conference on Machine Learning and Cybernetics (ICMLC), vol. 1, pp. 156-161. IEEE, 2017. [4]

5. Suhao, Li, Lin Jinzhao, Li Guoquan, Bai Tong, Wang Huiqian, and Pang Yu. "Vehicle type detection based on deep learning in traffic scene." Procedia computer science 131 (2018): 564-572. [5]

6. Li, Wei, Zhaowei Qu, Haiyu Song, Pengjie Wang, and Bo Xue. "The traffic scene understanding and prediction based on image captioning." IEEE Access 9 (2020): 1420-1427. [6]

7. Wojek, Christian, Stefan Walk, Stefan Roth, Konrad Schindler, and Bernt Schiele. "Monocular visual scene understanding: Understanding multi-object traffic scenes." IEEE transactions on pattern analysis and machine intelligence 35, no. 4 (2012): 882-897. [7]

8. Giachetti, Andrea, Marco Cappello, and Vincent Torre. "Dynamic segmentation of traffic scenes." In Proceedings of the Intelligent Vehicles' 95. Symposium, pp. 258-263. IEEE, 1995. [8]

9. S. Bhatlawande et al., "Study of Object Detection with Faster RCNN," 2022 2nd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2022, pp. 1-6. [9]

10. Song, Lei, Zheyuan Liu, Huixian Duan, and Na Liu. "A Color-Based Image Segmentation Approach for Traffic Scene Understanding." In 2017 13th International Conference on Semantics, Knowledge and Grids (SKG), pp. 33-37. IEEE, 2017. [10]

11. Kaggle Homepage, https://www.kaggle.com/, last accessed 2023/07/26. [11]

12. Chen, Long, Shaobo Lin, Xiankai Lu, Dongpu Cao, Hangbin Wu, Chi Guo, Chun Liu, and Fei-Yue Wang. "Deep neural network based vehicle and pedestrian detection for autonomous driving: A survey." IEEE Transactions on Intelligent Transportation Systems 22, no. 6 (2021): 3234-3246. [12]

13. Murugan, V., V. R. Vijaykumar, and A. Nidhila. "A deep learning RCNN approach for vehicle recognition in traffic surveillance system." In 2019 international conference on communication and signal processing (ICCSP), pp. 0157-0160. IEEE, 2019. [13]

14. Tariq, Abdullah, Muhammad Zeeshan Khan, and Muhammad Usman Ghani Khan. "Real time Vehicle Detection and Colour Recognition using tuned features of Faster-RCNN." In 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), pp. 262-267. IEEE, 2021. [14]

15. Oeljeklaus, Malte, Frank Hoffmann, and Torsten Bertram. "A fast multi-task CNN for spatial understanding of traffic scenes." In 2018 21st International Conference on Intelligent Transportation Systems (ITSC), pp. 2825-2830. IEEE, 2018. [15]

16. Deepajothi, S., D. Palanival Rajan, P. Karthikeyan, and S. Velliangiri. "Intelligent traffic management for emergency vehicles using convolutional neural network." In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), vol. 1, pp. 853-857. IEEE, 2021. [16]

THANK YOU

We would like to extend our sincere gratitude and thank The Reviewers, Organizing Committee Members, and Respected Session Chairs of the International Conference on Computing and Machine Learning (CML 2024) for giving us the opportunity to present our research paper.