IJACIS
International Journal of Advanced Computing and Intelligent System
MSIS

MSIS PRESS

Vehicle Detection System Through CNN

Md. Mehedi Hasan Naim, Md. Shafiul Islam Ayon, Md. Khaled Hasan, Md. Shohag Mia, Md. Khaled Hasan, Mr. Md. Abbas Ali Khan

mehedi15-14721@diu.edu.bd
shafiul15-14716@diu.edu.bd
khaled15-14691@diu.edu.bd
md15-14736@diu.edu.bd
abbas.cse@diu.edu.bd

Abstract: In this work, a reliable and effective model for the identification of cars in traffic surveillance. Vehicle detection is achieved through the Operation of Convolutional Neural Networks (CNNs), combining high-level and low-level information using feature fusion techniques. The use of static pictures is highly applicable for various operations in traffic surveillance. This makes it possible to identify different-sized cars using different attributes. We get 99.4% training accuracy & 99.3% in testing accuracy.

Keywords: 1. Autonomous Driving, 2.Real-time Vehicle Detection, 3. Target Tracking, 4. Continuous Projection of Geometry Features

1. Introduction

Vehicle detection is an essential component of modern transportation systems. It includes the automated detection of the presence, location and movement of vehicles inside an area using cutting-edge technology. Accurate and effective vehicle identification has drawn a lot of attention because of the fast rise of urbanization and the rising demands on traffic management, intelligent transportation systems, and autonomous cars[2].

The past several years have seen a substantial increase in vehicle identification systems due to advances in sensor technology, deep learning, and data accessibility. These systems employ a variety of technologies, such as image processing, machine learning, and neural networks, to recognize autos from many data sources, such as images, videos, and LiDAR scans. Researchers are investigating novel approaches to enhance the robustness and dependability of vehicle recognition systems in the face of challenges such as shifting weather patterns, occlusions, and intricate traffic situations.
Advancing vehicle detection techniques is pivotal for enhancing traffic management, enabling autonomous driving, and ensuring road safety in an increasingly automated and interconnected world[3][4]

2 Literature Review

In paper [1], it introduces YOLOv2_Vehicle, a novel vehicle detection model based on YOLOv2. It achieves a 94.78% mean Average Precision (mAP) on the BIT-Vehicle validation dataset and exhibits strong generalization on the CompCars test dataset.
The authors of [2] utilizes Haar-like features, AdaBoost algorithms, Gabor wavelet transform, and local binary pattern operator for vehicle detection and traffic surveillance. It achieves a detection rate over 97% with a false rate of only 3% and a vehicle recognition rate over 91%

In [3], the authors present YOLOv2_Vehicle.. Experimental results on the BIT-Vehicle validation dataset achieve a mean Average Precision (mAP) of 95%, showcasing its effectiveness.

Ronald Miller [4] surveys vision-based on-road vehicle detection systems for driver assistance.It focuses on camera-based systems mounted on vehicles rather than static monitoring setups.

A system developed in the United States named Autoscope, highlighting its advantages in paper [5].

The author in [6] proposes a hybrid DNN (HDNN) that divides maps into multiple blocks of variable sizes to

extract variable-scale features, outperforming traditional DNNs in vehicle detection.
In paper [7] explores vehicle detection and recognition in a traffic surveillance system using machine learning. It proposes the use of static image datasets and live CCTV surveillance, along with License Plate Recognition (LPR), for efficient vehicle verification and record-keeping.

Huaiyu Li [8], presents a vision-based vehicle detection and counting system for intelligent highway management using a new high definition dataset. The proposed methods demonstrate improved accuracy in detecting small vehicles and counting accurately.

The authors of paper [9] present a deep learning-based approach for accurate vehicle counting in traffic surveillance. Its experimental results demonstrate high accuracy (up to 99%).

3 Method

In methodology we ‘ll show our workflow diagram, Data Collection and Preprocessing Data Sources and Image Augmentation process where we collected two types of images & we do data augmentation on these data.

3.1 Workflow diagram

To identify a vehicle, we need to follow some effective steps. The workflow diaphragm of our proposed model is given bellow:

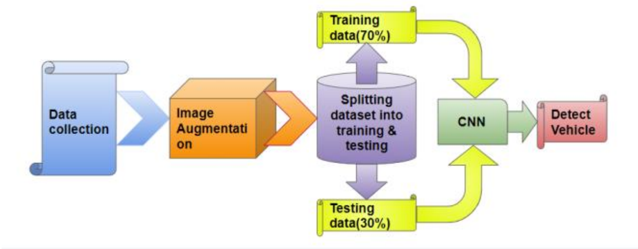

Figure 1: Workflow Diagram of Vehicle Detection Model

- Collect data & preprocess the input image.
- Extract features from the preprocessed image using Convolutional neural network (CNN).
- For each feature map, apply a sliding window technique to identify each window whether it is a vehicle or not.
- Filter out the false positives and select the regions with the highest probability of containing a vehicle. Then, Get the list of detected vehicles in output.

3.2 Data Collection and Preprocessing Data Sources

Collect a diverse dataset containing images and videos captured in various traffic scenarios, including urban and suburban areas.
Our dataset contains 2 parts.
These parts are:

1. Vehicle
2. Non-vehicle

$a_1 + a_2 + a_3 + \cdots + a_n = n$
(n=Total number of vehicle,aai=ith image)
$\sum_{i=1}^{n} a_i = n$——————————(1)

Figure 2: **Sample images of dataset**

Type	**Image**
Vehicle	
Non-vehicle	

b=Non-vehicle where

$b_1 + b + b_3 + \cdots + b_m = m$

(m=Total number of vehicle, b_i= i-th image)

$\sum_{i=1}^{m} b_i = m$——-----------------**(2)**

By using the addition technique in equation (1) and equation (2):

$\sum_{i=1}^{n} a_i + \sum_{i=1}^{m} b_i = n+m$.........(3)

=> vehicles + non-vehicles = (n+m)

Table 1: Parts of datas

3.3 Image Augmentation

Image augmentation is a technique which is widely used in computer vision and machine learning(ML) to increase the diversity of a training dataset by applying various transformations to existing images by using rotation,flip,scaling,translation,shearing,zooming techniques. The purpose of image augmentation is to increase the generalization ability of a machine learning(ML) model.

n rotation technique, we rotate the original image θ degree. $R_i = (Di, \theta)$

Where,

R_i=Rotated image, D_i=Original/Input image, θ=Angle of rotation

Here, the Original image is as input image & θ is the angle of rotation in degrees.

In Scaling technique, $S_i = resize(D_i, s_f)$

where, S_i=Scaled image, D_i=Original/Input image, s_i=scale factor

Here, scale factor is the factor by which the image is scaled. By using shifting technique in our 2D image set, $T_i = translate(D_i, \left(x_{shift}, y_{shift}\right))$

where,T_i=Translated image, D_i=Original/Input image, x_{shift}=shifting distance in x axis, y_{shift}=shifting distance in y axis

By flipping horizontally or vertically we can get a new image through data augmentation.

F_i=flip(D_i,axis=1) Where, F_i=Flipped image, D_i=Original/Input image

In zooming technique, $Z_i = zoom(D_i, z_f)$

where, Z_i=zoomed image, D_i=Original/Input image, Z_f=zoom factor. Here, the zoom factor determines the level of zooming.

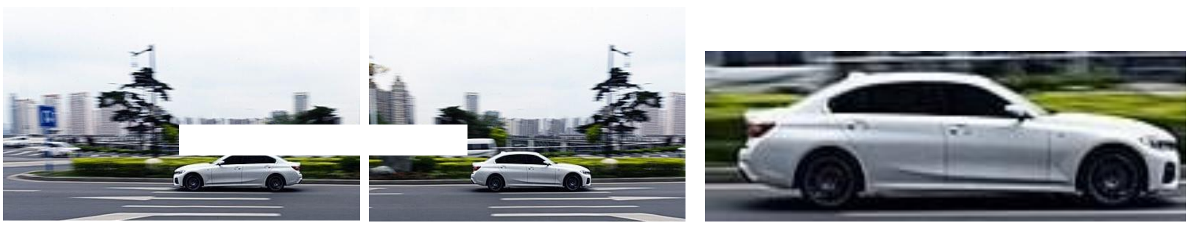

Figure 3: Real image & augmented images

Here, we used image augmentation technique where:

N=total number of vehicle, M=total number of non-vehicle

$\sum_{i=1}^{n}(a_i * k) = N$——-----------(4)

$\sum_{i=1}^{m}(b_i * k) = M$--------------

(5)

By using addition technique in equation (4) and equation (5):

$$\sum_{i=1}^{n}(a_i * k) + \sum_{i=1}^{m}(b_i * k) = N + M$$

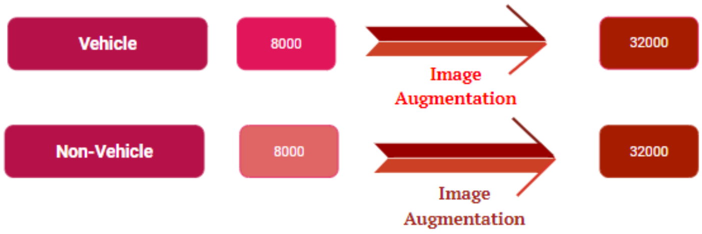

Figure 4: Data augmentation

Here, N=8000, M=3200, m=8000, M=3200 in my dataset. Here, k is the image augmentation multiple. Here,

$$8000 * k + 8000 * k = 64000$$

=> k = 4

So, we can tell that, after image augmentation our dataset increases .
Original Images *4 = Total number of images after image augmentation

Imageset details is:

Type	Train data	Test data	Total train data	Total test data
Vehicle	25600	6400	513200	12800
Non-vehicle	25600	6400		

Table 2: Dataset analysis

3.4 Training and testing

In this section, we ‘ll introduce the details of network training and testing using CNN.
We trained 70% of our data & tested 30% of it.

3.5 CNN Model for Feature Extraction Feature

Extraction: Implement a Convolutional Neural Network (CNN) to extract relevant features from the detected regions of interest (ROIs) within the images.
Architecture Selection: Design a CNN architecture suitable for extracting discriminative features specific to bikes, cycles, and rickshaws.
Training: Train the CNN on a dataset containing cropped ROIs of the detected vehicles. Ensure that the CNN learns
to distinguish between the target vehicle classes.

ReLU (Rectified Linear Units) is the real nonlinear function.

ReLU(x)=max(0,x)

ReLU works as an activation function, the correction layer replaces all the negative values by zeros. First Conv-1 is applied, the first set of convolutional filters to extract local features. In layer 2, Conv-2 - the second set of convolutional filters to extract features. Conv-3 used in layer 3, the third set of convolutional filters to capture high-level features.After that used Conv-4, the fourth set of convolutional filters to refine the features.Then, Conv-5 applies the fifth set of convolutional filters for further refinement. In layer 6, FC-6(Fully connected layer) that connects all the neurons from the previous layer. FC-7 is used after that. Fully connected layer that connects all the neurons from the previous layer.At the end, FC-8 is used as the final fully connected layer that produces the output prediction.

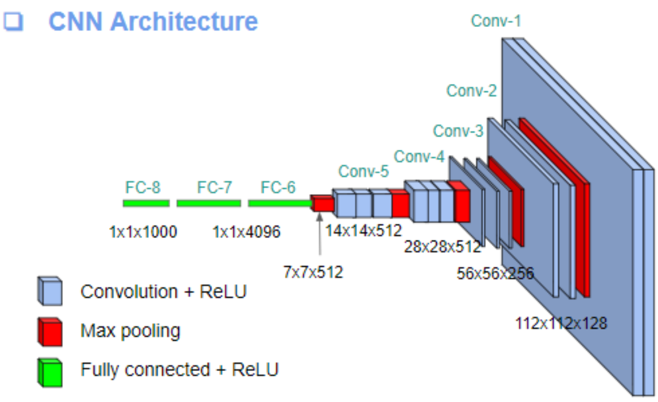

Figure 5: CNN Architecture

4. RESULTS AND DISCUSSION

Thorough training and testing on a carefully selected dataset were used to assess the effectiveness of the suggested CNN model for vehicle detection. The model's efficacy is demonstrated by the subsequent critical metrics:

4.1 Accuracy of Training

With a remarkable accuracy of 99.4%, the CNN performed exceptionally well during the training phase. The model's capacity to capture complex aspects relevant to vehicle recognition is demonstrated by its high training accuracy, which implies that it has effectively mastered the underlying patterns and features in the training dataset.

4.2 Accuracy of Testing

By means of extensive testing on a dataset that had not been encountered before, the CNN model's resilience was further evaluated. The remarkable 99.3% testing accuracy demonstrated the model's strong generalization to novel and unobserved cases. For the model to be practically used in real-world circumstances, where its dependability in recognizing automobiles is critical, this high testing accuracy is essential.

4.3 Precision, Recall, and F1-Score:

To offer a thorough view of the model's performance, several crucial metrics were computed in addition to accuracy, including precision, recall, and F1-score. Insights into the trade-offs between true positives, false positives, and false negatives are provided by these measures, providing a more comprehensive assessment.

	precision	recall	f1-score	support
non-vehicles	0.99	0.99	0.99	897
vehicles	0.99	0.99	0.99	879
accuracy			0.99	1776
macro avg	0.99	0.99	0.99	1776
weighted avg	0.99	0.99	0.99	1776

Figure 6: Precision, recall, F1-score, support

Precision=True positives/(True positives + False positives) ………..(1)

Recall=True positives/(True positives + False negatives)...................(2)

F1 score= 2* (precision*recall)/(precision+recall)............................(3)

4.4 Training & validation loss Versus training & validation accuracy

From our research work, we find the best epoch for train & validation loss is epoch 5. On the other hard, for train & validation accuracy we get epoch 2 as the best epoch.

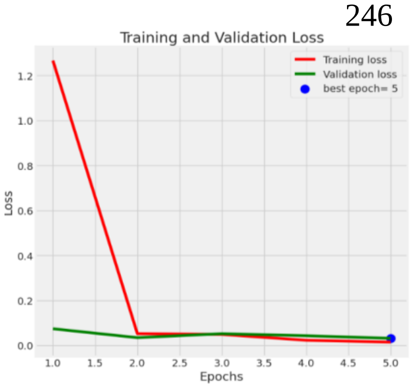

Figure 8.1: Training & validation loss

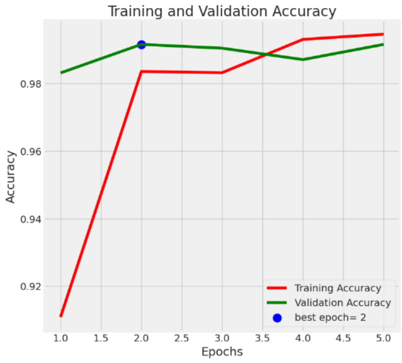

Figure 8.2: Training & Validation Accuracy.

4.5 Confusion Matrix:

A thorough examination of the confusion matrix, which breaks down the right and wrong classifications, clarifies the model's performance even further. Gaining insight into the model's behavior in various courses and comprehending possible areas for growth are made possible by this knowledge, which is priceless.

Here, $Accuracy = (tp + tn)/(tp + tn + fp + fn)$

tp=True positive, tn=True negation, fp=False positive, fn=False negative

Error= (1- Accuracy)

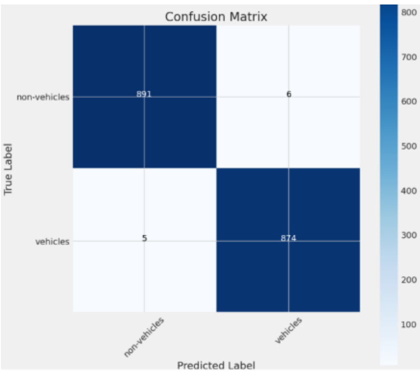

Figure 9: Confusion Matrix

4.6 Result comparison graph:

In this graph we can see that accuracy of our work is better than other research works. From this, we can say that we added some value in research regarding vehicle detection using CNN.

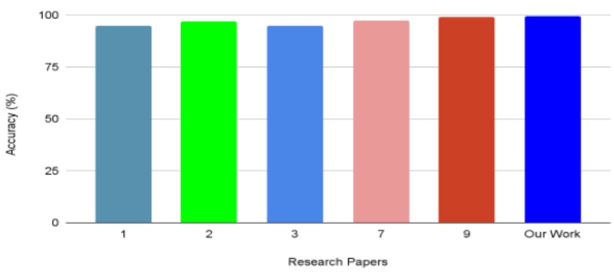

Figure 10: Result Comparison Graph

5. CONCLUSION

Our CNN-based vehicle detection model achieves excellent accuracy (99.3%) and works quickly in difficult situations, opening doors for advancements in transportation, surveillance, and autonomous driving applications and so on.

Future work in CNN-based vehicle detection involves making the model better at handling situations where vehicles are partially hidden, improving its speed, protecting it against malicious attacks, ensuring it can work well in different environments and properly integrating it with autonomous systems. Additionally, further research can focus on developing larger and more diverse datasets to enhance the model's ability to generalize and accurately detect vehicles in various real-world scenarios.

REFERENCES:

1. Yong Tang & Congzhe Zhang & Renshu Gu & Peng Li & Bin Yang (2015) Vehicle detection and recognition for intelligent traffic surveillance systems. Springer Science+Business Media New York 15 February 2015. DOI: 10.1007/s11042-015-2520-x **[Terjomefa]**

2. Jun Sang, Zhongyuan Wu, Pei Guo, Haibo Hu, Hong Xiang, Qian Zhang and Bin Cai (2018) An Improved YOLOv2 for Vehicle Detection **[MDPI]**

3, Zehang Sun, George B. and Miller (2004) On-Road Vehicle Detection Using Optical Sensors: A Review. 2004 IEEE Intelligent Transportation Systems Conference Washington, D.C., USA, October 36.2004 **[Researchgate]]**

4. Amirali Jazayeri, Cai (Member IEEE), Jiang Yu Zheng(Senior Member, IEEE) and Mihran Tuceryan(Senior Member, IEEE). Vehicle Detection and Tracking in Car Video Based on Motion Model. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 12, NO. 2, JUNE 2011 **[Researchgate]**

5. Xueyun Chen, Shiming Xiang, Cheng-Lin Liu, and Chun-Hong Pan. Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks. IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 11, NO. 10, OCTOBER 2014. **[ieee xplore]**

6. Joshua Gleason, Ara V. Nefian, Xavier Bouyssounousse, Terry Fong and George Bebis . Vehicle Detection from Aerial Imagery . 2011 IEEE International Conference on Robotics and Automation Shanghai International Conference Center May 9-13, 2011, Shanghai, China **[Academia]**

7. Sriashika Addala, Vehicle Detection and Recognition, May 2020, DOI:10.13140/RG.2.2.34908.82561.**[Researchgate]**

8. Mukesh Prasad , Chih-Ling Liu , Dong-Lin Li , Chandan Jha , Chin-Teng Lin, Multi-view Vehicle Detection based on Part Model with Active Learning, 2018 IEEE, **[IEEE]**

9. Haojia Lin, Zhilu Yuan2, Biao He, Xi Kuai 2, Xiaoming Li 2 and Renzhong Guo, A Deep Learning Framework for Video-Based Vehicle Counting, published: 21 February 2022, doi: 10.3389/fphy.2022.829734,[**frontiersin]**

10. Arthur, D.; Vassilvitskii, S. k-means++: The advantages of careful seeding. Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, New Orleans, LA, USA, 7–9 January 2007; pp. 1027–1035. **[Google Scholar]**

11. Neubeck, A.; Van G., L. Efficient non-maximum suppression. In Proceedings of the International Conference on Pattern Recognition (ICPR), Hong Kong, China, 20–24 August 2006 [**Google Scholar**]

12. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2005[**Google Scholar**]

13. Dong, Z.; Wu, Y.; Pei, M.; Jia, Y. Vehicle type classification using a semi supervised convolutional neural network. *IEEE Trans. Intel. Transp. Syst.* **2015** [**Google Scholar**] [**CrossRef**]

14. Yang, L.; Luo, P.; Change Loy, C.; Tang, X. A large-scale car dataset for fine-grained categorization and verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 3973–3981. [**Google Scholar**]

15. Zeiler, M.D.; Fergus, R. Visualizing and understanding convolutional networks. In Proceedings of the ECCV, Zurich, Switzerland, 6–12 September 2014; pp. 818–833. **[NYU]**

16. Ellis TJ. Rosin PL,Image Difference Threshold Strategies and Shadow Detection. Citeseer: BMVC (1995). p. 347–56.[**Researchgate**]

17. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 22–25 July 2017; pp. 2261–2269. [**Google Scholar**]

18. Pyo, J.; Bang, J.; Jeong, Y. Front collision warning based on vehicle detection using CNN. In Proceedings of the International SoC Design Conference (ISOCC), Jeju, Korea, 23–26 October 2016; pp. 163–164. [**Google Scholar**]

19. Tang, Y.; Zhang, C.; Gu, R. Vehicle detection and recognition for intelligent traffic surveillance system. *Multimed. Tools Appl.* **2017**, *76*, 5817–5832. [**Google Scholar**] [**CrossRef**]

20. Gao, Y.; Guo, S.; Huang, K.; Chen, J.; Gong, Q.; Zou, Y.; Bai, T.; Overett, G. Scale optimization for full-image-CNN vehicle detection. In Proceedings of the IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, USA, 11–14 June 2017; pp. 785–791. [**Google Scholar**]