Bangladesh Army University of Science and Technology (BAUST)		
Department of Computer Science and Engineering		
Assignment #1, Winter 2023	Level-4	Term-I
Course Code: CSE 4131	Course Title: Artificial Neural Networks and Fuzzy Systems	
Submission Date: 11-05-2023	CO Number: CO2	Full Marks: 15
ID: 200101037	Name: Aznur Azam	

Vehicle Detection Using CNN Models

1. Introduction

Vehicle detection is an important task in computer vision with various practical applications such as traffic monitoring, autonomous driving, and security systems. In recent years, deep learning techniques have achieved improved performance in vehicle detection tasks. Convolutional Neural Networks (CNNs) are particularly effective at learning features from images and making predictions based on those features. Vehicle detection using CNN models is a widely used method in computer vision and autonomous driving. CNNs are effective in analyzing visual data and excel at recognizing complex patterns. The goal is to identify and locate vehicles within images or video streams. The process involves dataset preparation, CNN architecture selection, model training, data augmentation, model evaluation, post-processing, and real-time inference. Vehicle detection using CNNs has applications in traffic monitoring, driver assistance systems, and autonomous vehicles, enabling safer and more efficient transportation systems. Vehicle detection using Convolutional Neural Network (CNN) models is a popular approach in computer vision and autonomous driving applications. CNNs are deep learning models specifically designed to analyze visual data, making them well-suited for tasks like object detection and classification.

2. Literature Review

2.1 Method 1

This paper proposes a method for detecting front vehicles in challenging weather conditions for autonomous cars. The method combines multifeatured fusion, dimensionality reduction, and CNN-based cascade detection. The approach achieves high recall rates of 98.69% and demonstrates good robustness with a recall rate of 97.32% in complex driving scenarios. Based on some data collected from BDD [39], Udacity, and the network, this paper selected 1200 pieces of data in good driving environments and complex driving environments with the size of 1280 × 720 to train the classifier and 600 pieces each to test the vehicle detection classifier. Moreover, in order to verify the performance of this algorithm, the detection results of multifeatured-fusion algorithm and the current mainstream detection algorithm are compared. The provided information does not explicitly mention any limitations or shortcomings of the method. To identify limitations, we would need more specific details or a more comprehensive description of the method. Without further information, it is difficult to determine the exact limitations of the cascade vehicle detection algorithm based on CNN.

2.2 Method 2

The paper proposes a new algorithm for detecting rear-approaching vehicles in agricultural machinery systems. It combines the Faster R-CNN deep learning model with structural similarity and root mean square comparison methods. This two-step approach achieves a detection rate of 98.2% and reduces false positives compared to general deep learning methods. The algorithm proves effective in accurately detecting rear-approaching vehicles in agricultural machinery systems. The paper presents a new algorithm for detecting rear-approaching vehicles in agricultural machinery systems. By combining deep learning and frame similarity techniques, the algorithm improves detection accuracy and reduces false positives. The algorithm utilizes the Faster R-CNN ResNet model for training, with a category set to 1 for vehicles and a learning step of 20,000. To address false positives, the algorithm compares the region of interest in frame images using the structural similarity (SSIM) index and mean square error (MSE). By analyzing the similarity values, the algorithm determines whether a vehicle is detected or not. The experimental results demonstrate that the algorithm successfully reduces false alarms caused by parked vehicles and passing vehicles captured by the rear camera. Overall, the algorithm offers an effective solution for detecting rear-approaching vehicles in agricultural machinery systems, improving safety for cultivator drivers.

The limitations of the proposed algorithm are as follows:

1. Lack of detailed analysis on false positive reduction.
2. Insufficient information on the reduction of unnecessary vehicle detections.
3. Limited dataset without information on size, diversity, and coverage of different scenarios.
4. Absence of discussion on performance and scalability on different hardware configurations or real-time implementation.
5. Lack of direct comparison with other deep learning models or algorithms for rear-approaching vehicle detection.
6. Addressing these limitations would provide a more comprehensive evaluation of the algorithm's performance and effectiveness.

2.3 Method 3

The paper introduces an automated system using convolutional neural networks (CNNs) for vehicle detection and counting in aerial images. It addresses challenges such as small vehicle size, diverse orientations, visual similarity with other objects, and road markings. The system generates a vehicle spatial density map and has been evaluated on the Munich and Overhead Imagery Research datasets. Experimental results show higher precision and recall rates compared to existing methods. Overall, the proposed system offers an efficient and effective approach for vehicle detection and counting in aerial images, with practical applications in urban planning and traffic management. The evaluation compared the proposed system with several state-of-the-art methods, including the Aggregated Channel Features (ACF) detector, ACF with Fast R-CNN, and Selective Search (SS) with Fast R-CNN. The performance metrics used were recall rate, precision rate, and F1-score. The proposed vehicle detection system using convolutional regression neural network has the following limitations:
1. Slow inference time compared to other existing methods, which may limit its applicability in real-time applications.
2. Subjectivity and inconsistency in detection results due to reliance on an empirical threshold for converting density maps into binary images.
3. Limited accuracy in complex or overlapping vehicle instances, leading to inaccurate counting or incomplete detection.
4. Lack of evaluation in diverse scenarios with varying lighting conditions, occlusions, or unusual vehicle shapes, which questions the system's generalizability and robustness.

2.4 Method 4

The paper introduces DP-SSD, a real-time vehicle detection framework for urban traffic surveillance. It enhances the Single Shot MultiBox Detector (SSD) by incorporating different feature extractors, deconvolution, and pooling operations. The framework adjusts the default box scale and achieves efficient vehicle detection on UA-DETRAC and KITTI datasets. With input sizes of 300 × 300 and 512 × 512, DP-SSD achieves mAP values of 75.43% at 50.47 FPS and 77.94% at 25.12 FPS, respectively, on an NVIDIA GeForce GTX 1080Ti GPU. DP-SSD performs on par with other state-of-the-art models, except for YOLOv3, which exhibits slightly better accuracy. The final models, DP-SSD300 and DP-SSD512, were tested on the KITTI test set consisting of 7518 images. The results, as shown in Table 7, demonstrate that DP-SSD512 achieves 85.32% mAP on the moderate difficulty level, surpassing R-SSD512 by 0.61%. Similarly, DP-SSD300 achieves a relatively high mAP of 83.86%, outperforming R-SSD300. Both DP-SSD models exhibit better accuracy than some existing methods. However, there is still a significant performance gap compared to the best method (e.g., THU CV-AI) reported on the KITTI website. Nevertheless, DP-SSD demonstrates relatively fast inference speeds. Figure 7 presents qualitative detection examples on the UA-DETRAC and KITTI test sets using the DP-SSD300 model, highlighting the robustness of the proposed method in detecting vehicles of different categories for urban traffic surveillance. One limitation of the paper is the removal of the batch normalization layer due to GPU memory limitations. Batch normalization is known to speed up network training and improve model accuracy, so its absence may have affected the performance of the DP-SSD framework.

Result and Analysis

1.The paper presents a multi-feature fusion algorithm for vehicle detection and compares its performance with other methods under good and complex weather conditions. The algorithm achieves a precision of 97.81% and an error rate of 2.19% under good weather conditions, and a precision of 95.73% and an error rate of 4.27% under complex weather conditions. The algorithm outperforms the compared methods in terms of accuracy and stability. However, in challenging weather conditions, misjudgments can occur due to similar textures and contour features. Overall, the multi-feature fusion algorithm demonstrates superior accuracy and stability in vehicle detection.

2. The proposed algorithm for rear-approach vehicle detection achieved high precision rates of 98.5% and 91.0% in Video 1 and Video 2, respectively. It effectively detected rear-approaching vehicles while minimizing false negatives. Additionally, a post-processing step utilizing frame similarities using the SSIM index and MSE was introduced to reduce false alarms caused by parked and passing vehicles. This further improved the algorithm's accuracy and performance. Although specific details of the results and analysis in Table 2 and Figure 6 were not provided, they likely provide further evidence of the algorithm's effectiveness in detecting rear-approaching vehicles and reducing false alarms.

3. the proposed system in this study achieved significant improvements in vehicle detection performance compared to state-of-the-art methods. It outperformed various detection algorithms on the Munich dataset, demonstrating superior precision, recall, and F1-score. The system also accurately detected vehicles in the OIRDS dataset, achieving high precision and recall rates. It showed low false positive rates and demonstrated effectiveness in challenging scenarios with occluded vehicles. Additionally, the system performed best on the scale it was trained on, but its performance decreased significantly with large changes in resolution. Overall, the proposed system showcased accurate vehicle detection, low false positive rates, and high precision and recall rates in aerial image datasets.

4. The researchers evaluated their DP-SSD model for vehicle localization and classification on UA-DETRAC and KITTI datasets. DP-SSD achieved mean average precision (mAP) scores of 75.43% and 77.94% for input sizes of 300x300 and 512x512, respectively, on the UA-DETRAC dataset. It also demonstrated real-time performance with frame rates of around 50.47 FPS and 25.12 FPS for the respective input sizes. Ablation experiments showed that DP-SSD outperformed traditional SSD models, thanks to enhanced feature pyramids and the simultaneous integration of deconvolution and pooling. Overall, DP-SSD proved effective for accurate vehicle localization and classification while maintaining real-time capabilities.

3. Conclusion

In conclusion, the researchers evaluated the DP-SSD model for vehicle localization and classification on UA-DETRAC and KITTI datasets. DP-SSD achieved high mean average precision (mAP) scores of 75.43% and 77.94% for different input sizes, demonstrating its accuracy. It also showed real-time performance with frame rates of around 50.47 FPS and 25.12 FPS. Ablation experiments confirmed that DP-SSD outperformed traditional SSD models by integrating enhanced feature pyramids and simultaneous deconvolution and pooling. The proposed algorithm achieved a 98.2% accuracy improvement compared to conventional deep learning methods, reducing false positive detections. Future work includes comparing with other deep learning models and utilizing frame motion and correlation for further reducing false positives. Additionally, a novel vehicle detection and counting method using convolutional regression neural network was introduced, outperforming state-of-the-art methods in terms of true positive rate, false alarm rate, F1 score, and precision. However, the proposed system requires more inference time, and future work aims to develop a faster model with improved performance.

References

1. Private Car Ownership in China in 2019." AskCI, 7 January 2020, http//www.askci.com/news/chanye/20200107/1624461156140.shtml.

2. Chen, C. Rear Approaching Vehicle Detection with Microphone. Bachelor's Thesis, Halmstad University, 2013.

3. Zheng, Z., Wang, X., Zhou, G., and Jiang, L. "Vehicle Detection Based on Morphology from Highway Aerial Images." Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, July 2012, pp. 5997-6000.

4. Chu, W., Liu, Y., Shen, C., Cai, D., and Hua, X.S. "Multi-Task Vehicle Detection With Region-of-Interest Voting." IEEE Transactions on Image Processing, vol. 27, 2018, pp. 432-441.

5. Hu, X., Xu, X., Xiao, Y., Chen, H., He, S., Qin, J., and Heng, P.A. "SINet: A Scale-insensitive Convolutional Neural Network for Fast Vehicle Detection." arXiv, 2018, arXiv:1804.00433v1.

6. Kang, H.W., Baek, J.W., and Jeong, Y.S. "Real-Time Side-Rear Vehicle Detection Algorithm for Blind Spot Warning Systems." KIISE Transactions on Computing Practice, vol. 23, no. 7, 2018, pp. 408-416.

7. Batavia, P.H., Pomerleau, D.A., and Thorpe, C.E. "Overtaking Vehicle Detection Using Implicit Optical Flow." Proceedings of the Conference on Intelligent Transportation Systems, Boston, MA, USA, 12 November 1997.

8. Maji, S., Berg, A.C., and Malik, J. "Classification using intersection kernel support vector machines is efficient." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2008, pp. 1-8.