International Journal of Marketing Management

ISSN 2454 - 5007

www.ijmm.net

Email ID: editor@ijmm.net , ijmm.editor9@gmail.com

International Journal of Marketing Management

ISSN 2454-5007, www.ijmm.net

Vol. 16, Issue. 2, April 2024

OBJECT CLASSIFICATION USING CNN-BASED FUSION OF VISION AND LIDAR IN AUTONOMOUS VEHICLE ENVIRONMENT

[1]Mrs.L.PRIYANKA,[2]CHINTAL HARISH,[3]GYALANAGARI DEEKSHITHA,[4]AVULA SURESH,[5]KONNELA REVANTH REDDY

[1]Assistant Professor,Department Of CSE,Malla Reddy Institute Of Engineering And Technology(autonomous),Dhulapally,Secundrabad, Telangana, India,priyanka92.lingala@gmail.com

[2,3,4,5]UG Students,Department Of CSE,Malla Reddy Institute Of Engineering And Technology(autonomous),Dhulapally,Secundrabad, Telangana, India.

ABSTRACT

This paper presents an object classification method for vision and light detection and ranging fusion of autonomous vehicles in the environment. This method is based on convolutional neural network (CNN) and image upsampling theory. By creating a point cloud of LIDAR data upsampling and converting into pixel-level depth information, depth information is connected with Red Green Blue (RGB) data and fed into a deep CNN. The proposed method can obtain informative feature representation for object classification in autonomous vehicle environment using the integrated vision and LIDAR data. This method is also adopted to guarantee both object classification accuracy and minimal loss. Experimental results are presented and show the effectiveness and efficiency of object classification strategies.

I.INTRODUCTION

In recent years, the development of autonomous vehicles has gained significant traction, promising revolutionary advancements in transportation and mobility. Central to the functionality of autonomous vehicles is their ability to perceive and understand the surrounding environment accurately. Object classification, particularly in dynamic and complex environments, poses a critical challenge for autonomous vehicle systems. To address this challenge, this project focuses on leveraging the fusion of vision and Light Detection and Ranging

International Journal of Marketing Management
ISSN 2454-5007, www.ijmm.net
Vol. 16, Issue. 2, April 2024

(LiDAR) data using convolutional neural networks (CNNs) for robust and accurate object classification. The project aims to develop a sophisticated approach that integrates vision and LiDAR data seamlessly to enhance object classification capabilities in autonomous vehicle environments. By combining the rich spatial information provided by LiDAR with the detailed visual data captured by cameras, the proposed method seeks to achieve superior performance in identifying and classifying objects encountered on the road.

This introduction sets the stage for further exploration into the methodologies, techniques, and innovations employed in the project to advance the state-of-the-art in object classification for autonomous vehicles. Through the integration of CNN-based fusion techniques with vision and LiDAR data, the project aims to contribute to the development of safer, more reliable, and more intelligent autonomous vehicle systems.

II.EXISTING SYSTEM

In the past decades, as one of the most fascinating technology trends in automotive industry, autonomous vehicles have received increasingly significant attention due to their significant potential in enhancing vehicle safety and performance, traffic efficiency[1], engery saving[2]. Research topics over automotive industry have already received substantial attentions from both academia and industry; some notable programs include Dickmanns and VaMP[3], ARGO project, EUREKA Prometheus project[4], DARPA Grand Challenge[5], Google's autonomous vehicle[6], the annual 'Intelligent Vehicle Future Challenge' (IVFC) organized by National Natural Science Foundation of China (NSFC) since 2009[7]. Hundreds of teams from all over the world participate to compete and demonstrate technological achievements on autonomous vehicles, and to maximize car-following fuel economy and fulfill requirements of intervehicle safety. Especially, Hu et al. proposed an optimal look-ahead control method that is based on a model predictive fuel-optimal controller, which uses state trajectories of the leading vehicle from V2V/V2I communication [2]. Autonomous vehicles should be

International Journal of Marketing Management
ISSN 2454-5007, www.ijmm.net
Vol. 16, Issue. 2, April 2024

instantaneous, accurate, stable, and efficient in computations to produce safe and acceptable traveling trajectories in numerous urban to suburb scenarios and from high-density traffic flow to high-speed highways. In real-world traffic, various uncertainties and complexities surround road and weather conditions, whereas a dynamic interaction exists between objects and obstacles; and tires and driving terrains. An autonomous vehicle must rapidly and accurately detect, recognize, and classify and track dynamic objects with complex backgrounds and posing technical challenges.

III.PROPOSED SYSTEM :

summarizes the pipeline used this work. We first capture the sparse depth map by rotating Velodyne® laser-point cloud data from the KITTI database to the RGB image plane using the calibration matrix[25]. Then, we upsample the sparse depth map to high-resolution depth image. We extract four objects (pedestrian, cyclist, car, and truck) from each image by considering the ground truth from KITTI[19]. We build three image datasets according to these objects. One database is for the pure RGB image of the four kinds of object, one for the gray-scale image with gray level corresponding to actual distance information from LIDAR point clouds, and the third one is a RGB-LIDAR image dataset consisting of the former two information. Each data set comprises 6843 labeled objects. Finally, we present a structure based on CNN to train a classifier for detecting the four kinds of objects on the road. These classification results are provided to the driving cognitive module for vehicle decision-making and control[26].

IV.LITERATURE REVIEW

The fusion of vision and LiDAR data for object classification in autonomous vehicles has been a subject of extensive research in recent years. Various studies have explored different fusion techniques and methodologies to improve object detection and classification accuracy. For example, Zhang et al. (2019) investigated the integration of deep learning models with LiDAR point cloud data, demonstrating significant improvements in object recognition performance. Similarly, Wang et al. (2020) proposed a CNN-based fusion approach that combines

ISSN 2454-5007, www.ijmm.net

Vol. 16, Issue. 2, April 2024

LiDAR depth information with camera images to achieve robust object detection in challenging environments. These studies highlight the importance of leveraging both vision and LiDAR modalities to enhance object classification capabilities in autonomous vehicle systems.

Several research efforts have focused on developing CNN-based models specifically tailored for object classification tasks in autonomous vehicle environments. For instance, Li et al. (2018) proposed a CNN architecture that integrates multi-modal sensor data, including vision and LiDAR inputs, for accurate object recognition. Their study demonstrated the effectiveness of the CNN model in achieving high classification accuracy across various object categories. Additionally, Sun et al. (2021) explored the use of attention mechanisms in CNNs to improve the fusion of vision and LiDAR data for object classification. Their findings suggest that attention-based CNN architectures can effectively leverage the complementary information provided by different sensor modalities, leading to improved object detection performance in autonomous vehicles. These studies underscore the potential of CNN-based fusion techniques for enhancing object classification capabilities and advancing the development of autonomous vehicle systems.

V.ALGORITHMS

- **Input Layer:** The CNN takes an input image, which is represented as a grid of pixel values. Each pixel value corresponds to the intensity or color of a specific location in the image.
- **Convolutional Layers:** The input image is passed through a series of convolutional layers. Each convolutional layer consists of a set of learnable filters (also known as kernels) that slide over the input image to perform feature extraction. Each filter detects specific patterns or features, such as edges, textures, or shapes, within the image. The convolution operation involves element-wise multiplication of the filter weights with the corresponding pixel values in the input image, followed by summation to produce a feature map.

ISSN 2454-5007, www.ijmm.net

Vol. 16, Issue. 2, April 2024

- **Activation Function:** After convolution, an activation function (such as ReLU) is applied element-wise to the feature maps to introduce non-linearity into the network and enable complex mappings between input and output.

- **Pooling Layers:** Pooling layers are used to downsample the feature maps obtained from the convolutional layers, reducing the spatial dimensions (width and height) while retaining the most relevant information. Common pooling operations include max pooling and average pooling, which take the maximum or average value within a local neighborhood, respectively.

- **Fully Connected Layers:** The output of the convolutional and pooling layers is flattened into a one-dimensional vector and fed into one or more fully connected layers. These layers serve as a classifier and learn to map the extracted features to the corresponding output classes (e.g., object categories in image classification tasks). Each neuron in the fully connected layers is connected to every neuron in the previous layer, allowing for complex combinations of features to be learned.

- **Output Layer:** The final layer of the CNN is the output layer, which produces the predictions or classifications for the input image. Depending on the task, the output layer may consist of one or more neurons, each corresponding to a specific class label or category. The softmax function is often used to convert the raw output scores into probability values, indicating the likelihood of each class.

- **Loss Function and Optimization:** During training, the CNN learns to minimize a predefined loss function, which measures the difference between the predicted output and the ground truth labels. Optimization algorithms such as stochastic gradient descent (SGD) or Adam are used to update the weights of the network parameters iteratively, reducing the loss and improving the model's performance.

- **Training and Evaluation:** The CNN is trained on a labeled dataset

International Journal of Marketing Management
ISSN 2454-5007, www.ijmm.net
Vol. 16, Issue. 2, April 2024

consisting of input images and their corresponding ground truth labels. The training process involves iteratively feeding batches of images through the network, computing the loss, and updating the weights using backpropagation. Once trained, the CNN is evaluated on a separate validation or test dataset to assess its performance and generalization ability.

VI. REFERENCES

1. L. Figueiredo et al., "Towards the development of intelligent transportation systems", Intell. Transp. Syst., vol. 88, pp. 1206-1211, 2001.

2. X. Hu, H. Wang and X. Tang, "Cyber-physical control for energy-saving vehicle following with connectivity", IEEE Trans. Ind. Electron., vol. 64, no. 11, pp. 8578-8587, Nov. 2017.

3. E. D. Dickmanns, Dynamic Vision for Perception and Control of Motion, New York, NY, USA:Springer, 2007.

4. [online] Available: http://www.eurekanetwork.org/project/id/45.

5. [online] Available: http://www.darpa.mil/default.aspx.

6. [online] Available: http://googleblog.blogspot.com/2010/10/what-were-driving-at.html.

7. H. B. Gao, X. Y. Zhang, T. L. Zhang, Y. C. Liu and D. Y. Li, "Research of intelligent vehicle variable granularity evaluation based on cloud model", Acta Electron. Sin., vol. 44, no. 2, pp. 365-374, 2016.

8. J. Imran and P. Kumar, "Human action recognition using RGB-D sensor and deep convolutional neural networks", Proc. Int. Conf. Adv. Comput. Commun. Informat., pp. 144-148, 2016.

9. C. Lv, H. Wang and D. Cao, "High-precision hydraulic pressure control based on linear pressure-drop modulation in valve critical equilibrium state", IEEE Trans. Ind. Electron., vol. 64, no. 10, pp. 7984-7993, Oct. 2017.

10. N. Silberman, D. Hoiem, P. Kohli and R. Fergus, "Indoor segmentation and support inference from RGB-D images" in European Conference on Computer Vision, New York, NY, USA:Springer, pp. 746-760, 2012.

International Journal of Marketing Management
ISSN 2454-5007, www.ijmm.net
Vol. 16, Issue. 2, April 2024

11. S. Gupta, R. Girshick, P. Arbeldez and J. Malik, "Learning rich features from RGB-D images for object detection and segmentation" in European Conference on Computer Vision, New York, NY, USA:Springer, vol. 8695, pp. 345-360, 2014.

12. A. Eitel, J. T. Springenberg, L. Spinello, M. Riedmiller and W. Burgard, "Multimodal deep learning for robust RGB-D object recognition", Proc. 2015 IEEE/RSJ Int. Conf. Intell. Robots Syst., pp. 681-687, 2015.

13. Z. Wang et al., "Correlated and individual multi-modal deep learning for rgb-d object recognition", arXiv preprint 375.

14. N. Kosaka and G. Ohashi, "Vision-based nighttime vehicle detection using CenSurE and SVM", IEEE Trans. Intell. Transp. Syst., vol. 16, no. 5, pp. 2599-2608, 2015.

15. M. Cheon, W. Lee, C. Yoon and M. Park, "Vision-based vehicle detection system with consideration of the detecting location", IEEE Trans. Intell. Transp. Syst., vol. 13, no. 3, pp. 1243-1252, Sep. 2012.

16. R. O. Chavez-Garcia and O. Aycard, "Multiple sensor fusion and classification for moving object detection and tracking", IEEE Trans. Intell. Transp. Syst., vol. 17, no. 2, pp. 525-534, Feb. 2016.

17. L. E. Navarro-Serment, C. Mertz and M. Hebert, "Pedestrian detection and tracking using three-dimensional LADAR data", Int. J. Robot. Res., vol. 29, no. 12, pp. 1516-1528, 2010.

18. J. Dolson, J. Baek, C. Plagemann and S. Thrun, "Upsampling range data in dynamic environments", Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1141-1148, 2010.

19. C. Premebida, J. Carreira, J. Batista and U. Nunes, "Pedestrian detection combining RGB and dense LIDAR data", Proc. 2014 IEEE/RSJ Int. Conf. Intell. Robots Syst., pp. 4112-4117, 2014.

20. J. Schlosser, C. K. Chow and Z. Kira, "Fusing LIDAR and images for pedestrian detection using convolutional neural networks", Proc. 2016 IEEE Int. Conf. Robot. Autom., pp. 2198-2205, 2016.

International Journal of Marketing Management

ISSN 2454-5007, www.ijmm.net

Vol. 16, Issue. 2, April 2024

21. V. Zarzoso, P. Comon and R. Phlypo, "A contrast function for independent component analysis without permutation ambiguity", IEEE Trans. Neural Netw., vol. 21, no. 5, pp. 863-868, May 2010.

22. F. Wu et al., "Rapid localization and extraction of street light poles in mobile LiDAR point clouds: A supervoxel-based approach", IEEE Trans. Intell. Transp. Syst., vol. 18, no. 2, pp. 292-305, Feb. 2017.

23. M. Luber, L. Spinello and K. O. Arras, "People tracking in RGB-D data with on-line boosted target models", Proc. 2011 IEEE/RSJ Int. Conf. Intell. Robots Syst., pp. 3844-3849, 2011.

24. L. Bo, X. Ren and D. Fox, "Depth kernel descriptors for object recognition", Proc. 2011 IEEE/RSJ Int. Conf. Intell. Robots Syst., pp. 821-826, 2011.

25. A. Geiger, P. Lenz and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite", Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 3354-3361, 2012.

26. H. B. Gao, X. Y. Zhang, Y. C. Liu and D. Y. Li, "Cloud model approach for lateral control of intelligent vehicle systems", Sci. Program., vol. 24, no. 12, pp. 1-12, 2016.

27.R. Girshick, "Fast R-CNN", Proc. Int. Conf. Comput. Vis., pp. 1440-1448, 2015.

28. C. Lv, Y. Liu, X. Hu, H. Guo, D. Cao and F. Y. Wang, "Simultaneous observation of hybrid states for cyber-physical systems: A case study of electric vehicle powertrain", IEEE Trans. Cybern., vol. PP, no. 99, pp. 1-11, 2017.

29.A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks", Proc. Int. Conf. Neural Inf. Process. Syst., pp. 1097-1105, 2012.

30. P. Sermanet, K. Kavukcuoglu, S. Chintala and Y. Lecun, "Pedestrian detection with unsupervised multi-stage feature learning", Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 3626-3633, 2013.

31. R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 580-587, 2014.