In: 2024 IEEE International Workshop on Metrology for Automotive (MetroAutomotive), 26 - 28 Jun 2024, Bologna, Italy, pp. 41-46
DOI:10.1109/metroautomotive61329.2024.10615661

Enhancing Object Detection and Localization through Multi-Sensor Fusion for Smart City Infrastructure

Soujanya Syamal
Centre for Autonomous and Cyber-Physical System, SATM
Cranfield University
Bedford, United Kingdom
soujanya.syamal@cranfield.ac.uk

Cheng Huang
Centre for Autonomous and Cyber-Physical System, SATM
Cranfield University
Bedford, United Kingdom
cheng-huang.huang@cranfield.ac.uk

Ivan Petrunin
Centre for Autonomous and Cyber-Physical System, SATM
Cranfield University
Bedford, United Kingdom
i.petrunin@cranfield.ac.uk

Abstract— **The rapid advancement in autonomous systems and smart city infrastructure demands sophisticated object detection and localization capabilities to ensure safety, efficiency, and reliability. Traditional single sensor approaches often fall short, especially under complex environmental conditions. This paper introduces the CLR-Localiser, a novel multi-sensor fusion framework that synergistically integrates data from cameras, LiDAR and radar sensors mounted on roadside infrastructure to enhance object detection and 3D localization. Leveraging the complementary strengths of each sensor type, the CLR-Localiser employs an early fusion approach and deep learning techniques, including convolutional neural networks for object detection and regression networks for precise localization. We rigorously validated the performance of the CLR-Localiser against the benchmark Kitti dataset, and a custom dataset specifically designed for this research, demonstrating significant improvements in detection accuracy, localization precision, and object-tracking capabilities under diverse conditions. Our findings highlight the CLR-Localiser's potential to overcome the limitations of conventional monocular and single-sensor methods, offering a robust solution for autonomous driving, robotics, surveillance, and industrial automation applications. The development and validation of the CLR-Localiser not only prove the technical feasibility of early sensor data fusion but also pave the way for future advancements in multi-sensor fusion technology for enhanced environmental perception in autonomous systems.**

Keywords— Multi-Sensor Fusion, Object Detection, 3D Localization, Autonomous Systems, Deep Learning, LiDAR, Radar, Camera, Environmental Perception, Smart City Infrastructure

I. INTRODUCTION

This research focuses on enhancing environmental perception for autonomous systems and advanced robotics by developing a multi-sensor fusion framework. Traditional object detection methods, such as those using monocular cameras, struggle with poor lighting, occlusions, and unclear object representations. By integrating data from Radar, LiDAR, and Cameras through deep learning, our framework aims to overcome these challenges. This multi-sensor approach is crucial for improving object localization and detection in autonomous driving, robotics, surveillance, and industrial automation, where accuracy and reliability are key.
The goal is to create a robust and adaptable fusion framework that leverages the unique advantages of each sensor type to enhance object detection and localization across various real-world conditions.
The main contributions of this paper are as follows:

(a) A robust multi-sensor fusion framework integrating camera, radar, and LiDAR data to enhance object detection and localization. This framework optimizes the use of each sensor, improving accuracy and reliability in diverse environments.
(b) We introduce novel deep learning architectures for sensor data fusion, designed to integrate heterogeneous data from cameras, radar, and LiDAR. This approach enhances environmental perception and autonomous decision-making, representing a significant advancement in multi-sensor fusion technology.
(c) Development of a comprehensive custom dataset for training and evaluating multi-sensor systems, covering various scenarios. This dataset provides unparalleled flexibility, enabling precise tuning and optimization of our system over general datasets by its specialized nature and the proprietary methodologies involved.
We conducted extensive testing of our novel architecture on both our custom dataset and the widely recognized Kitti dataset, with results closely aligning across both. This parallel performance validates our architecture's effectiveness and its potential to stand alongside state-of-the-art techniques in the field.
These advancements aim to significantly benefit autonomous systems and intelligent robotics, promoting safer and more efficient applications.

II. RELATED WORKS

In the evolving domain of multi-sensor fusion for object detection and localization, leveraging Radar, LiDAR, and camera data through deep learning has been at the forefront of research, aiming to address the inherent limitations of single-sensor systems in complex environments.
The works of [1] and [2] have significantly advanced the utilization of Radar for object detection in adverse weather conditions, underscoring its ability to complement visual sensors under conditions of poor visibility. Additionally, the critical role of LiDAR technology in achieving precise object localization in cluttered settings has been demonstrated by some contributions [3], [4], highlighting its indispensable high-resolution 3D mapping capabilities. The importance of cameras for capturing detailed visual features, despite their sensitivity to variable lighting, was emphasized in the research by [5], illustrating their critical role in comprehensive environmental perception.
The paradigm shift to deep learning for object detection and localization was marked by the adoption of Convolutional Neural Networks (CNNs) as the standard. Foundational

insights provided by [6], [7] underscore the superiority of CNNs in learning hierarchical features, facilitating this transition. The development and refinement of meta-architectures such as Faster R-CNN, R-FCN, and SSD were critically evaluated by [8], [9], [10], [11], each contributing to the balance between accuracy, computational efficiency, and speed in object detection tasks.

The integration of heterogeneous sensor data has been explored through various fusion strategies. Early statistical fusion techniques by [12], [13] laid the groundwork, which was expanded upon by [14] through the introduction of deep learning-based methods capable of handling complex data dimensions. The works of [15], [16], [17], [18], [19] further propelled the field by developing architectures for real-time sensor data fusion, highlighting the benefits of combining different sensor modalities to overcome environmental challenges.

The exploration of uncertainty in sensor data fusion through Stochastic Neural Networks (SNNs) and Bayesian Neural Networks (BNNs) has opened new avenues for research. Studies by [20], [21], [22], [23] delved into these models, comparing their robustness and calibration capabilities. Particularly, [24] proposed a novel single-stage fusion approach using BNNs to integrate camera, LiDAR, and RADAR data, setting a new direction for future research in multi-sensor fusion.

Despite the progress, the field faces ongoing challenges related to data heterogeneity, sensor alignment, and computational demands. The works of [25], [26], [27] highlight these challenges while suggesting future research directions, including the exploration of new sensor modalities and advanced deep learning techniques to enhance fusion strategies and improve system performance.

Most existing studies, such as those by [1], [2], [3], [4], [5], have focused on the capabilities of individual sensors or the integration of two sensor types, often overlooking the comprehensive potential of fusing Radar, LiDAR, and camera data simultaneously. Furthermore, while studies like [15], [16], [18], [19] have explored deep learning methods for sensor fusion, there is a conspicuous absence of a unified framework that can effectively handle the heterogeneity of sensor data, align features from different sensor types, and optimize computational efficiency for real-time applications. Also, in smart city applications some research can be seen where sensor fusion by deep learning for running autonomous vehicle framework has been described [28].

Building upon the extensive literature review, it's evident that while significant advancements have been made in multi-sensor fusion for object detection and localization, there remains a critical gap in integrating Radar, LiDAR, and camera data through a unified, deep learning-based framework that can operate efficiently in real-time across diverse environmental conditions. This research addresses this gap by proposing a novel multi-sensor fusion framework that not only leverages the unique strengths of each sensor modality but also introduces advanced deep learning techniques to enhance the accuracy and robustness of object detection and localization tasks.

III. METHODOLOGY

This research introduces an advanced multi-sensor fusion architecture as depicted in Figure 1, the CLR-Localiser, designed to enhance object detection and localization by leveraging the synergistic potential of camera, LiDAR and radar data. The system's efficacy is grounded in its ability to accurately identify and track objects across diverse environmental conditions, using deep learning techniques and the Deep SORT algorithm for dynamic object trajectory management.

A. *Multi-Sensor Fusion Architecture*

At the core of our methodology is a novel multi-sensor fusion architecture shown is Figure 1 that integrates data at the input level, facilitating early fusion. This approach ensures that the heterogeneous data from the three sensors is synchronized and calibrated [29], aligning them within a common spatial and temporal framework. This fusion strategy is pivotal in harnessing the complementary strengths of each sensor type: the depth and precision of LiDAR, the rich visual details provided by cameras, and the motion detection capabilities of radar.

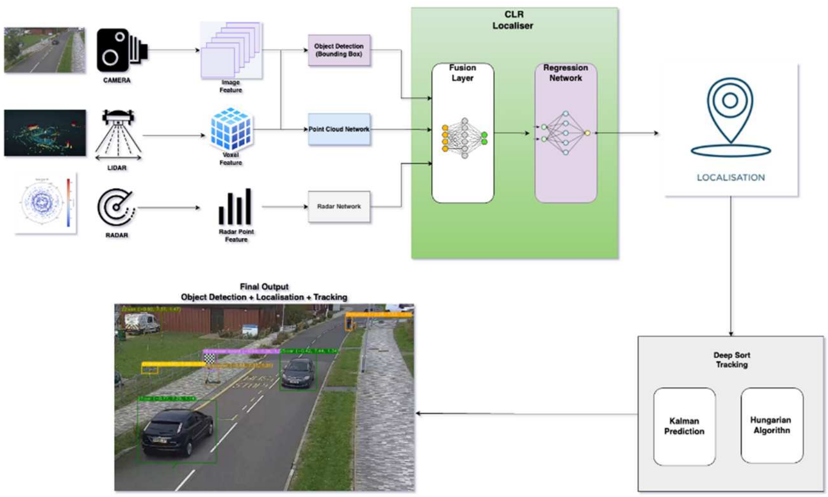

Figure 1: The High-Level Architecture

B. Data Acquisition and Synchronization

Data from the three sensors are meticulously synchronized, ensuring temporal and spatial coherence. The camera serves as the primary sensor for initial object detection, employing a convolutional neural network (CNN) trained on a custom dataset. LiDAR data augments this by offering precise 3D localization, while radar data introduces additional motion context. The synchronization process ensures that data from all sensors corresponds to the same environmental snapshot, enabling accurate fusion and analysis.

C. Object Detection and Localization

The CLR-Localiser employs a deep learning model for object detection, utilizing camera data to produce bounding boxes for detected objects. LiDAR data is processed to extract precise spatial coordinates, enhancing localization accuracy. Radar data is then fused to refine object motion and positioning further. This staged process culminates in a comprehensive data set that accurately reflects the objects' positions and movements within the environment.

D. Fusion and Regression Network

A key innovation in our methodology is the development of a regression-based network architecture tailored for multi-modal data processing. This network combines inputs from the camera, LiDAR, and radar, applying deep learning techniques to compute accurate localization metrics. The fusion network outputs refined localization for each detected object, significantly improving upon the precision achievable with single-sensor methods.

E. Object Tracking

Object tracking is achieved through the integration of the Deep SORT algorithm [30], The Deep SORT algorithm significantly enhances object tracking by integrating deep learning with classical tracking techniques. It extends the SORT algorithm by employing deep neural networks for feature extraction, allowing for robust tracking across frames by effectively handling occlusions and appearance changes. Deep SORT utilizes a Kalman filter for motion prediction and the Hungarian algorithm for object assignment, improving tracking accuracy and reducing identity switches in dynamic environments. Its application in multi-sensor fusion scenarios further demonstrates its versatility and contribution to enhancing the environmental perception of autonomous systems.

F. Evaluation Metrics

The performance of the CLR-Localiser is rigorously evaluated against established metrics, including Mean Squared Error (MSE) for localization accuracy and Intersection over Union (IoU) for detection precision. Additional metrics, such as precision, recall, and F1 score, further quantify the system's efficacy in object detection and tracking tasks.

This methodology represents a significant advancement in multi-sensor fusion technology, offering a robust framework for the precise detection and localization of objects in complex environments. By effectively combining and analyzing data from LiDAR, cameras, and radar, the CLR-Localiser sets a new benchmark for accuracy and reliability in autonomous systems and intelligent robotics.

IV. RESULTS

Our study evaluated the CLR-Localiser, an innovative system for integrating multiple sensors (Camera, LiDAR, Radar) to detect, locate, and track objects accurately in real scenarios. Tests were conducted at Cranfield University's MUEAVI Track, a facility dedicated to advancing autonomous technologies. The experimental setup featured an Ouster 64-beam LiDAR, PTZ camera, and 2-D MM wave radar, using TensorFlow and PyTorch for integration [31].

A. Object Detection Performance

Utilizing a custom dataset, our CNN model achieved remarkable detection performance, underscored by precision and recall metrics. The diversity and balance within our dataset were critical for model training, addressing class imbalance through targeted data augmentation strategies. The confusion matrix (Figure 2) and precision-recall and F1 Confidence curves (Figure 3 and Figure 4) further validated our model's accuracy, achieving a mean average precision (mAP) of 0.967 for object detection tasks.

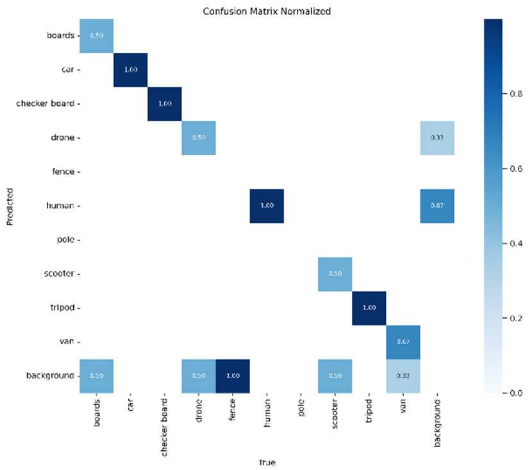

Figure 2: Confusion Matrix for Object Detection

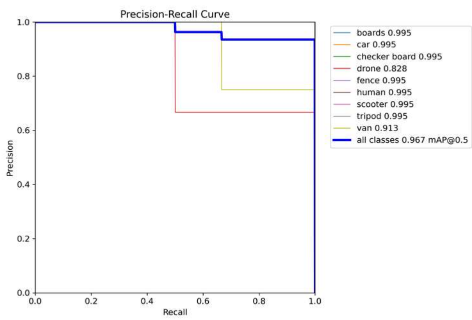

Figure 3: Precision-Recall Curve for Object Detection

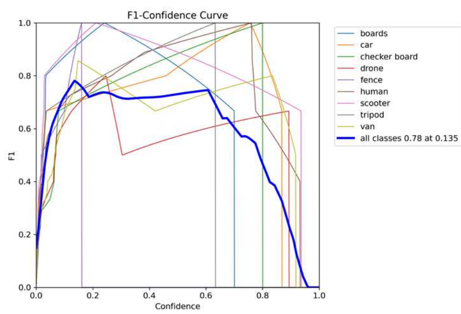

Figure 4: F1 confidence curve, for Object detection

B. Sensor Fusion and Localization Accuracy

The early fusion of LiDAR, camera, and radar data facilitated enhanced object localization. Figure 5 and Figure 6 illustrate the seamless Projection of sensor data, with LiDAR points projected into camera frames and radar information augmenting the detection process. This multi-modal approach significantly improved the robustness and precision of object localization, demonstrating the system's capability to leverage the strengths of each sensor modality effectively.

Figure 5: Projection of Lidar Data into Camera Frames

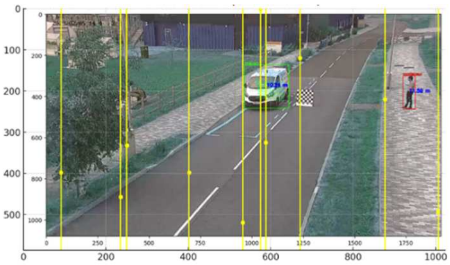

Figure 6: Projection of radar data into the Lidar and Camera Frame

C. Object Localization

Table 1: Localization metrics

Metric	Value
Localization Precision:	0.95
Localization Recall	0.94
Mean X-axis Error	0.0045 meters
Mean Y-axis Error	0.047 meters
Mean Z-axis Error	0.83 meters

The regression network within CLR-Localiser was rigorously evaluated for its localization accuracy. Results indicated a high degree of precision (0.95) and recall (0.94), with minimal mean axis errors, showcasing the network's ability to predict accurate 3D locations of detected objects. These metrics highlight the CLR-Localiser's competency in combining data from multiple sensors to achieve precise object localization. The results can be seen in Table 1 and Figure 7.

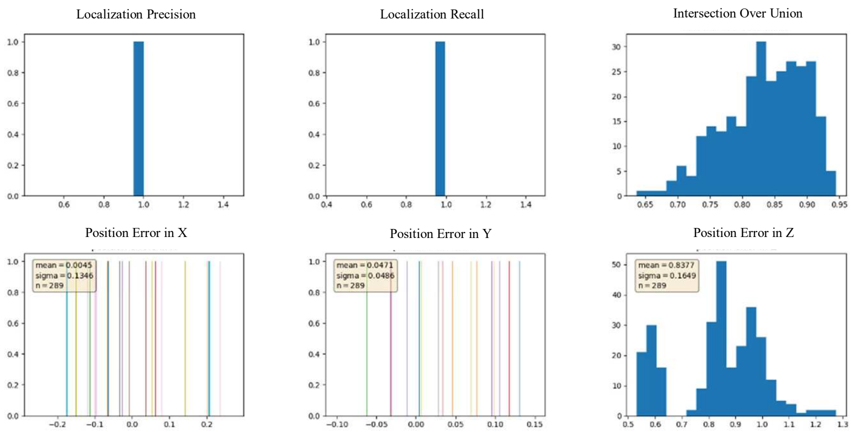

Figure 7 Localization Results with Errors

D. Object Tracking

Incorporating the Deep SORT algorithm, our system exhibited strong tracking capabilities. RMSE metrics for single and multiple objects reflected the system's reliability in maintaining accurate object trajectories over time as seen in Figure 8.

Table 2: Results of the tracking

Object Tracked	RMSE
Car (Big and Far)	0.12
Human (Big and Near)	0.15
Scooter (small and far)	0.19

Real-time testing, including both static and dynamic scenes, further affirmed the system's robustness in tracking, with precise localization metrics obtained for various object sizes and distances mentioned in Table 2.

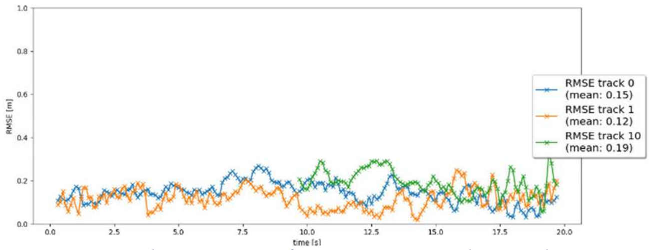

Figure 8: Tracking metrics of testing scenario showing low RMSE of the tracking system.

E. Real-world Applicability

Quantitative results from real-time testing underscored the CLR-Localiser's potential for deployment in practical applications. Achieving a localization precision of 0.85 and recall of 0.82, the system demonstrated its effectiveness across diverse real-world conditions. Figure 9 offers a visual testament to the system's tracking accuracy, showcasing its capability to handle complex scenarios with multiple, moving objects.

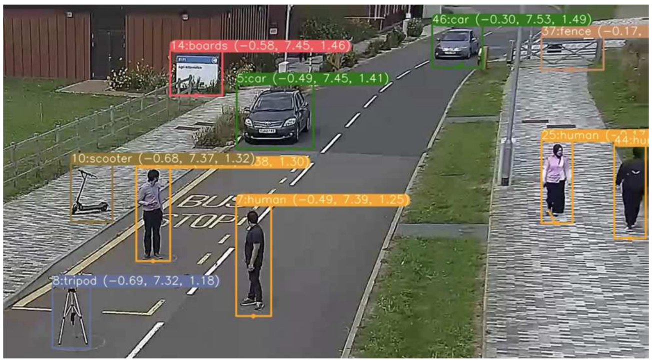

Figure 9: Testing the system in real-world scenario in MUEAVI Track of Cranfield University

The CLR-Localiser evaluation highlights its progress in multi-sensor fusion for precise object detection and tracking by using LiDAR, camera, and radar. It shows potential for autonomous systems and surveillance. Future efforts will aim to improve depth estimation accuracy to boost its real-world performance.

V. Discussions

The evaluation of our CLR-Localiser system, which integrates camera, LiDAR and radar for enhanced object detection and tracking, highlights its strengths and the challenges of multi-sensor fusion. This analysis connects our results with the wider field, moving beyond numbers to critically understand the implications and insights gained.

A. Object Detection Analysis

Our system shows strong object detection with high precision and recall, indicating reliable recognition. However, the need for a more diverse and balanced dataset is evident, as current accuracy varies across classes, raising questions about real-world applicability. Precision-recall trade-offs also highlight the need for caution in scenarios where false positives could have significant consequences.

B. Sensor Fusion Insights

The fusion approach enhances object localization by combining the strengths of various sensors, despite challenges like data synchronization and sensor alignment. Analysis identifies areas for improvement, particularly in LiDAR depth estimation and 2D radar capabilities. Validation using custom and Kitti dataset (Figure 10) shows a promising localization precision of 95.8% for pedestrians, with minimal errors of 0.003, 0.02 and 0.054 meters on the x, y, and z axes, but underscores the necessity for further real-world testing.

Table 3: MAP Comparison with state-of-the-art architecture

Dataset	Method	MAP
Ours (Custom)	CLR-Localiser	95%
Kitti	CLR-Localiser	95.8%
Waymo	Deep-Fusion	84.37%
nuScenes	CFR-net	55.99%

Also, some comparison has been done with state-of-the-art architecture such as Deep-Fusion [32] and CFR-net [33]. And the statistic on Table 3 shows the precision of CLR-localiser is better than the rest of the method compared.

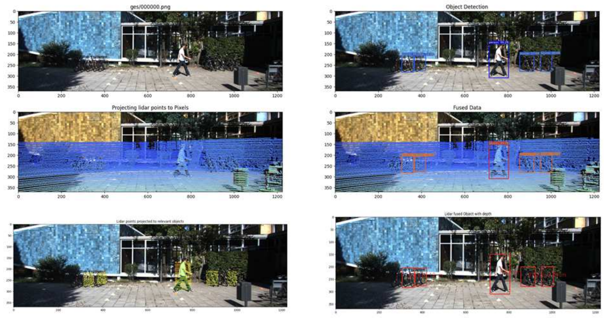

Figure 10: Validation with KITTI dataset

C. Localization Performance

The CLR-Localiser shows high localization accuracy, especially on the X and Y axes, but struggles with Z-axis (depth) errors due to limitation of 2D radar, highlighting the need for better depth perception and refined sensor fusion techniques for accurate distance measurements.

D. Real-world Tracking Evaluation

Real-world tracking (Figure 11) demonstrates the system's adaptability and quick response in dynamic settings, though tracking accuracy decreases with smaller object sizes, indicating a need for algorithm improvement. While performance is stable across different distances, further testing on scalability and robustness in varied conditions is needed.

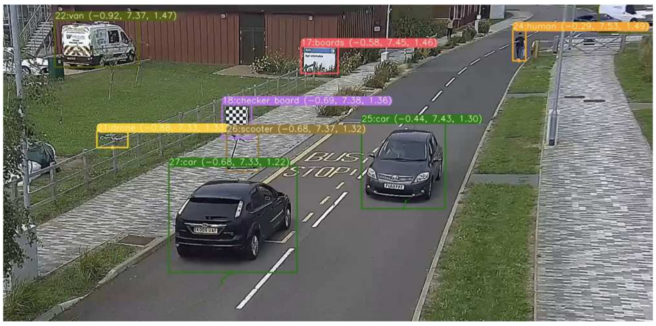

Figure 11: Tracking in Real World Scenario

E. Future Directions

Our analysis recognizes the CLR-Localiser's strengths and areas for improvement, like depth estimation and vertical localization. Its adaptability and real-world potential are promising, but challenges call for ongoing refinement. Future work will focus on advanced fusion techniques, refining small object tracking, and expanding real-world tests. This journey highlights the iterative process of technological progress.

VI. Conclusion

This study concludes the CLR-Localiser system's development, integrating camera, LiDAR, and radar for advanced object detection and tracking. Highlighting a novel multi-sensor fusion approach, our system demonstrates high accuracy and reliability in real-world settings. Despite its success, challenges like depth estimation and tracking small objects remain. Future work aims to enhance depth perception and expand testing. The CLR-Localiser marks a significant step forward in fusion technology, promising to impact autonomous systems and robotics.

REFERENCES

[1] M. Khalil *et al.*, "An Improved Approach for Multi-Target Detection and Tracking in Automotive Radar Systems".

[2] A. Manjunath, Y. Liu, B. Henriques, and A. Engstle, "Radar Based Object Detection and Tracking for Autonomous Driving".

[3] S. Muhammad and G. W. Kim, "Visual object detection based lidar point cloud classification," *Proceedings - 2020 IEEE International Conference on Big Data and Smart Computing, BigComp 2020*, pp. 438–440, Feb. 2020, doi: 10.1109/BIGCOMP48618.2020.00-32.

[4] Z. Xiang, "A Simultaneous Object Detection and Tracking Framework Based on Point Cloud," *2022 IEEE 14th International Conference on Computer Research and Development, ICCRD 2022*, pp. 225–229, 2022, doi: 10.1109/ICCRD54409.2022.9730356.

[5] P. Wang, "Research on comparison of LiDAR and camera in autonomous driving," *J Phys Conf Ser*, vol. 2093, no. 1, Nov. 2021, doi: 10.1088/1742-6596/2093/1/012032.

[6] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257–276, May 2019, doi: 10.1109/JPROC.2023.3238524.

[7] A. Mauri *et al.*, "Deep Learning for Real-Time 3D Multi-Object Detection, Localisation, and Tracking: Application to Smart Mobility," *Sensors 2020, Vol. 20, Page 532*, vol. 20, no. 2, p. 532, Jan. 2020, doi: 10.3390/S20020532.

[8] X. Wang, A. Shrivastava, and A. Gupta, "A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 3039–3048, Apr. 2017, doi: 10.1109/CVPR.2017.324.

[9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans Pattern Anal Mach Intell*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[10] A. Pfeuffer and K. Dietmayer, "Optimal Sensor Data Fusion Architecture for Object Detection in Adverse Weather Conditions," *2018 21st International Conference on Information Fusion, FUSION 2018*, pp. 2588–2595, Jul. 2018, doi: 10.23919/ICIF.2018.8455757.

[11] J. Huang *et al.*, "Speed/accuracy trade-offs for modern convolutional object detectors," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 3296–3305, Nov. 2016, doi: 10.1109/CVPR.2017.351.

[12] G. G. Cunto and J. Z. Sasiadek, "Sensor fusion ins/gnss based on fuzzy logic weighted kalman filter," *International Conference of Control, Dynamic Systems, and Robotics*, pp. 1–8, 2020, doi: 10.11159/CDSR20.129.

[13] P. Vangasse and D. Nicholson, "Architectures for Efficient Data Fusion," *Multisensor Fusion*, pp. 197–221, 2002, doi: 10.1007/978-94-010-0556-2_8.

[14] A. Pfeuffer and K. Dietmayer, "Optimal Sensor Data Fusion Architecture for Object Detection in Adverse Weather Conditions," *2018 21st International Conference on Information Fusion, FUSION 2018*, pp. 2588–2595, Sep. 2018, doi: 10.23919/ICIF.2018.8455757.

[15] J. Alfred Daniel, C. Chandru Vignesh, B. Anand Muthu, R. Senthil Kumar, C. Sivaparthipan, and C. Enrique Montenegro Marin, "Fully convolutional neural networks for LIDAR-camera fusion for pedestrian detection in autonomous vehicle", doi: 10.1007/s11042-023-14417-x.

[16] I. Ogunrinde and S. Bernadin, "Deep Camera–Radar Fusion with an Attention Framework for Autonomous Vehicle Vision in Foggy Weather Conditions," *Sensors 2023, Vol. 23, Page 6255*, vol. 23, no. 14, p. 6255, Jul. 2023, doi: 10.3390/S23146255.

[17] M. P. Muresan, M. Raul, S. Nedevschi, and R. Danescu, "Stereo and Mono Depth Estimation Fusion for an Improved and Fault Tolerant 3D Reconstruction," *Proceedings - 2021 IEEE 17th International Conference on Intelligent Computer Communication and Processing, ICCP 2021*, pp. 233–240, 2021, doi: 10.1109/ICCP53602.2021.9733702.

[18] A. Gonzalez, G. Villalonga, J. Xu, D. Vazquez, J. Amores, and A. M. Lopez, "Multiview random forest of local experts combining RGB and LIDAR data for pedestrian detection," *2015 IEEE Intelligent Vehicles Symposium (IV)*, vol. 2015-August, pp. 356–361, Aug. 2015, doi: 10.1109/IVS.2015.7225711.

[19] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, "Multi-view 3D object detection network for autonomous driving," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 6526–6534, Nov. 2017, doi: 10.1109/CVPR.2017.691.

[20] R. Ravindran, M. J. Santora, and M. M. Jamali, "Camera, LiDAR, and Radar Sensor Fusion Based on Bayesian Neural Network (CLR-BNN)," *IEEE Sens J*, vol. 22, no. 7, pp. 6964–6974, Apr. 2022, doi: 10.1109/JSEN.2022.3154980.

[21] J. Mitros and B. Mac Namee, "On the Validity of Bayesian Neural Networks for Uncertainty Estimation," *CEUR Workshop Proc*, vol. 2563, pp. 140–151, Dec. 2019, Accessed: Mar. 08, 2024. [Online]. Available: https://arxiv.org/abs/1912.01530v2

[22] Y. Ovadia *et al.*, "Can You Trust Your Model's Uncertainty? Evaluating Predictive Uncertainty Under Dataset Shift," *Adv Neural Inf Process Syst*, vol. 32, Jun. 2019, Accessed: Mar. 08, 2024. [Online]. Available: https://arxiv.org/abs/1906.02530v2

[23] A. Kristiadi, M. Hein, and P. Henni, "Being Bayesian, Even Just a Bit, Fixes Overconfidence in ReLU Networks," *37th International Conference on Machine Learning, ICML 2020*, vol. PartF168147-7, pp. 5392–5402, Feb. 2020, Accessed: Mar. 08, 2024. [Online]. Available: https://arxiv.org/abs/2002.10118v2

[24] R. Ravindran, M. J. Santora, and M. M. Jamali, "Camera, LiDAR, and Radar Sensor Fusion Based on Bayesian Neural Network (CLR-BNN)," *IEEE Sens J*, vol. 22, no. 7, pp. 6964–6974, Apr. 2022, doi: 10.1109/JSEN.2022.3154980.

[25] D. Gohring, M. Wang, M. Schnurmacher, and T. Ganjineh, "Radar/Lidar sensor fusion for car-following on highways," *ICARA 2011 - Proceedings of the 5th International Conference on Automation, Robotics and Applications*, pp. 407–412, 2011, doi: 10.1109/ICARA.2011.6144918.

[26] B. S. Jahromi, T. Tulabandhula, and S. Cetin, "Real-Time Hybrid Multi-Sensor Fusion Framework for Perception in Autonomous Vehicles," *Sensors 2019, Vol. 19, Page 4357*, vol. 19, no. 20, p. 4357, Oct. 2019, doi: 10.3390/S19204357.

[27] O. Mees, A. Eitel, and W. Burgard, "Choosing smartly: Adaptive multimodal fusion for object detection in changing environments," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2016-November, pp. 151–156, Nov. 2016, doi: 10.1109/IROS.2016.7759048.

[28] S. Syamal, J. Datta, S. Basu, and S. Das, "Deep Learning-Based Autonomous Driving and Cloud Traffic Management System for Smart City," *Convergence of IoT, Blockchain, and Computational Intelligence in Smart Cities*, pp. 117–135, Jan. 2023, doi: 10.1201/9781003353034-7/DEEP-LEARNING-BASED-AUTONOMOUS-DRIVING-CLOUD-TRAFFIC-MANAGEMENT-SYSTEM-SMART-CITY-SOUJANYA-SYAMAL-JOYATEE-DATTA-SRIJITA-BASU-SHUVENDU-DAS.

[29] J. Domhof, K. F. P. Julian, and K. M. Dariu, "An extrinsic calibration tool for radar, camera and lidar," *Proc IEEE Int Conf Robot Autom*, vol. 2019-May, pp. 8107–8113, May 2019, doi: 10.1109/ICRA.2019.8794186.

[30] R. Kejriwal, H. J. Ritika, A. Arora, and Mohana, "Vehicle Detection and Counting using Deep Learning basedYOLO and Deep SORT Algorithm for Urban Traffic Management System," *2022 1st International Conference on Electrical, Electronics, Information and Communication Technologies, ICEEICT 2022*, 2022, doi: 10.1109/ICEEICT53079.2022.9768653.

[31] Y. Li *et al.*, "DeepFusion: Lidar-Camera Deep Fusion for Multi-Modal 3D Object Detection," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2022-June, pp. 17161–17170, 2022, doi: 10.1109/CVPR52688.2022.01667.

[32] Y. Li *et al.*, "DeepFusion: Lidar-Camera Deep Fusion for Multi-Modal 3D Object Detection", Accessed: Mar. 15, 2024. [Online]. Available: https://github.com/NVIDIA/semantic-segmentation

[33] F. Nobis, M. Geisslinger, M. Weber, J. Betz, and M. Lienkamp, "A Deep Learning-based Radar and Camera Sensor Fusion Architecture for Object Detection", Accessed: Mar. 15, 2024. [Online]. Available: https://github.com/TUMFTM/CameraRadarFusionNet

Enhancing object detection and localization through multi-sensor fusion for smart city infrastructure

Syamal, Soujanya

2024-06-26

Syamal S, Huang C, Petrunin I. (2024) Enhancing object detection and localization through multi-sensor fusion for smart city infrastructure. In: 2024 IEEE International Workshop on Metrology for Automotive (MetroAutomotive), Proceeding Volume 32, 26-28 June 2024, Bologna, Italy, pp. 41-46

https://doi.org/10.1109/metroautomotive61329.2024.10615661