NIGHTTIME PEDESTRIAN DETECTION BASED ON A FUSION OF VISUAL INFORMATION AND MILLIMETER-WAVE RADAR

[1]MR.P.UDAY, [2]MAMILLAPALLY PARASHARAMULU, [3]RAJALINGOLLA GANGA SUNITH, [4]GUJJA SAI CHARAN REDDY, [5]SERU PRASHANTH RAJ, [6]KANDI BHARATH SAI

[1]Assistant Professor, Department of CSE, Malla Reddy Engineering College. Hyderabad, Telangana

[2,3,4,5,6]Students, Department of CSE, Malla Reddy Engineering College. Hyderabad, Telangana

ABSTRACT

Nighttime pedestrian detection is a critical component in modern intelligent transportation systems and autonomous driving, where visibility conditions are significantly degraded due to low illumination, shadows, and environmental noise. Traditional vision-based systems often struggle to accurately detect pedestrians at night due to limited lighting and poor contrast. To address these challenges, this project proposes a Nighttime Pedestrian Detection System based on the fusion of visual information and millimeter-wave radar, combining complementary sensing modalities to improve detection accuracy and robustness. The proposed system integrates infrared or low-light cameras with millimeter-wave radar sensors to capture both visual and spatial information. The visual subsystem extracts features such as shape, motion, and texture from images using deep learning models like Convolutional Neural Networks (CNNs), while the radar subsystem provides distance, velocity, and object localization data that are less affected by lighting conditions. A sensor fusion module is employed to combine these heterogeneous data sources using techniques such as feature-level fusion or decision-level fusion. This integration enhances the system's ability to detect pedestrians even in challenging environments such as fog, rain, or complete darkness. Experimental results demonstrate that the fusion-based approach significantly outperforms standalone vision or radar systems in terms of detection accuracy, precision, and reliability. The system shows improved performance in reducing false positives and false negatives, ensuring safer navigation in real-world scenarios. Additionally, the use of radar enables continuous detection regardless of lighting conditions, while visual data provides contextual information for better classification. This hybrid approach is particularly suitable for applications in autonomous vehicles, advanced driver assistance systems (ADAS), and smart surveillance systems. Overall, the proposed system contributes to enhancing road safety by providing a reliable and efficient solution for nighttime pedestrian detection.

Keywords: Nighttime Pedestrian Detection, Sensor Fusion, Millimeter-Wave Radar, Computer Vision, Deep Learning, Autonomous Vehicles, ADAS, Infrared Imaging, Object Detection, Intelligent Transportation Systems

I.INTRODUCTION

The advancement of intelligent transportation systems and autonomous driving technologies has increased the need for reliable pedestrian detection systems, especially under challenging nighttime conditions. During nighttime, visibility is significantly reduced due to poor illumination, shadows, and environmental noise, making it difficult for traditional vision-based systems to accurately detect pedestrians. Conventional object detection models such as YOLO and SSD have shown strong performance in daylight scenarios but often suffer from reduced accuracy in low-light environments [1], [6]. Infrared imaging techniques have been introduced to address this limitation; however, they may lack sufficient contextual information and are affected by environmental factors such as temperature variations [11], [13]. These challenges highlight the necessity for advanced detection systems capable of operating effectively under diverse conditions. Therefore, improving nighttime pedestrian detection has become a critical research area to enhance road safety and reduce accidents in real-world driving environments [10].

To overcome the limitations of single-sensor systems, recent research has focused on multi-sensor fusion techniques, combining visual data with complementary sensing technologies such as millimeter-wave radar. Radar sensors provide accurate distance, velocity, and object localization information regardless of lighting conditions, making them highly suitable for nighttime applications [20], [21]. Meanwhile, visual sensors capture rich semantic and contextual information necessary for object classification. Fusion techniques, including early fusion, feature-level fusion, and decision-level fusion, have demonstrated improved detection accuracy by leveraging the strengths of both modalities [23], [24]. Studies have shown that integrating radar and camera data significantly enhances detection performance in complex environments, including fog, rain, and low-light scenarios [22], [40]. However, challenges such as data alignment, synchronization, and computational complexity remain open research problems in multimodal systems.

In this context, the proposed project, Nighttime Pedestrian Detection Based on a Fusion of Visual Information and Millimeter-Wave Radar, aims to develop a robust and efficient detection framework by combining deep learning-based vision models with radar sensing capabilities. The system utilizes advanced object detection architectures such as YOLOv5 and its variants to extract visual features, while radar data provides spatial and motion information for improved localization [15], [18]. A fusion module integrates these heterogeneous data sources to enhance detection accuracy and reduce false positives. Additionally, techniques such as attention mechanisms and Kalman filtering are employed to improve tracking and stability [36], [38]. The proposed approach ensures reliable pedestrian detection under challenging nighttime conditions, contributing to safer autonomous driving and advanced driver assistance systems (ADAS).

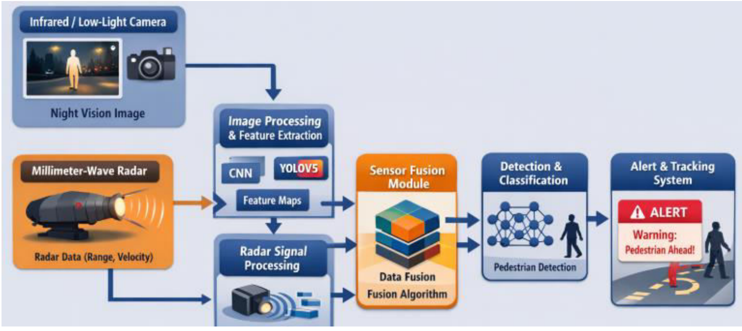

Figure1: System Architecture for Nighttime Pedestrian Detection Using Visual and Millimeter-Wave Radar Fusion

This figure represents the overall architecture of the proposed Nighttime Pedestrian Detection System, which integrates both visual sensing (infrared/low-light camera) and millimeter-wave radar to achieve accurate detection in low-visibility conditions. The system begins with two parallel input sources: the infrared camera, which captures night vision images, and the radar sensor, which provides distance, velocity, and spatial information independent of lighting conditions. The visual data is processed through an image processing and feature extraction module, typically using deep learning models such as CNN or YOLOv5 to generate feature maps. Simultaneously, radar signals undergo radar signal processing to extract meaningful features such as object position and motion. Both feature sets are then combined in a sensor fusion module, where advanced fusion algorithms integrate visual and radar information to enhance detection accuracy and robustness. The fused data is passed to the detection and classification module, which identifies pedestrians and distinguishes them from other objects. Finally, the alert and tracking system generates warnings (e.g., "Pedestrian Ahead") and tracks detected individuals in real time. This architecture ensures reliable performance under challenging nighttime conditions, improving safety in autonomous driving and advanced driver assistance systems (ADAS).

II SURVEY OF RESEARCH

The work by J. Redmon et al. (2016) [1] introduced the YOLO (You Only Look Once) framework, a real-time object detection system that revolutionized computer vision by performing detection in a single forward pass. The methodology uses a unified neural network that divides the image into grids and predicts bounding boxes and class probabilities simultaneously. The results demonstrated high speed and competitive accuracy, making YOLO suitable for real-time applications such as autonomous driving. However, the early versions of YOLO struggled with small object detection and performance under low-light conditions. This limitation is critical in nighttime pedestrian detection scenarios. Nevertheless, this study forms the foundation for modern detection systems, and improved versions of YOLO are utilized in the proposed system for efficient feature extraction.

J. Redmon and A. Farhadi (2017, 2018) [2], [3] further improved YOLO through YOLO9000 and YOLOv3, enhancing detection accuracy and introducing multi-scale predictions. The methodology incorporated deeper network architectures and anchor boxes to improve detection of objects at different scales. The results showed significant improvements in accuracy and robustness compared to earlier versions. However, challenges remained in detecting objects in low-light or noisy environments. These studies highlight the importance of deep learning-based object detection, which is extended in the proposed system using enhanced models such as YOLOv5 for nighttime applications.

W. Liu et al. (2016) [6] proposed the SSD (Single Shot MultiBox Detector), which performs object detection using multi-scale feature maps and convolutional neural networks. The methodology allows detection at different resolutions, improving performance for objects of varying sizes. The results demonstrated a good balance between speed and accuracy. However, SSD also faces challenges in detecting small and distant objects, particularly under poor lighting conditions. This limitation emphasizes the need for additional sensing modalities, such as radar, to improve detection reliability in nighttime environments.

S. M. Patole et al. (2017) [10] provided a comprehensive review of automotive radar systems, focusing on signal processing techniques used for object detection and tracking. The methodology highlights how radar can accurately measure distance, velocity, and angle of objects regardless of lighting conditions. The results show that radar is highly reliable in adverse environments such as fog, rain, and darkness. However, radar lacks detailed visual information required for object classification. This study supports the integration of radar with vision systems, forming the basis of the sensor fusion approach used in the proposed system.

J. Kim et al. (2020) [20] proposed a sensor fusion network combining radar and camera data for 3D object detection. The methodology uses radar range-azimuth maps along with image features to improve detection accuracy. The results demonstrate that fusion-based models outperform single-sensor approaches in complex driving scenarios. However, challenges such as synchronization and alignment of heterogeneous data remain. This research strongly supports the use of fusion techniques in the proposed system to enhance nighttime pedestrian detection performance.

Y. Zhou et al. (2022) [40] presented a comprehensive review of millimeter-wave radar and camera fusion technologies in intelligent transportation systems. The study discusses various fusion strategies, including early, middle, and late fusion techniques. The results highlight that fusion systems significantly improve detection accuracy and robustness under challenging environmental conditions. However, the study also identifies issues such as computational complexity and data integration challenges. This research validates the effectiveness of multimodal fusion and supports the design of the proposed system, which combines visual and radar data for reliable nighttime pedestrian detection.

III. WORKING METHODOLOGY

The proposed Nighttime Pedestrian Detection System based on Visual and Millimeter-Wave Radar Fusion follows a multi-stage methodology to ensure accurate and reliable detection under low-light conditions. The process begins with data acquisition, where inputs are collected from two primary sensors: an infrared or low-light camera and a millimeter-wave radar. The camera captures visual information such as pedestrian shape, texture, and movement, while the radar provides spatial data including distance, velocity, and object location, which is unaffected by lighting conditions. In the next stage, data preprocessing and feature extraction are performed separately for both modalities. The visual data is processed using deep learning models such as YOLOv5 or CNN-based architectures, which extract feature maps representing objects in the scene. Simultaneously, radar signals undergo signal processing techniques to extract meaningful features such as range, Doppler velocity, and angle of arrival. These features are then structured into radar feature maps for further analysis. Noise reduction and normalization techniques are applied to both data streams to improve quality and consistency. The extracted features are then integrated using a sensor fusion module, which combines visual and radar data through techniques such as feature-level fusion or decision-level fusion. This fusion enhances detection accuracy by leveraging the complementary strengths of both sensors—visual data for classification and radar data for localization. The fused data is then passed to the detection and classification module, where the system identifies pedestrians and differentiates them from other objects. Finally, the system includes a tracking and alert mechanism, where detected pedestrians are continuously tracked using algorithms such as Kalman filtering. If a pedestrian is detected within a critical range, the system generates a real-time alert to ensure safety. This methodology ensures high accuracy, robustness, and real-time performance, making it suitable for applications in autonomous vehicles and advanced driver assistance systems (ADAS).

IV RESULTS EXPLANATIONS

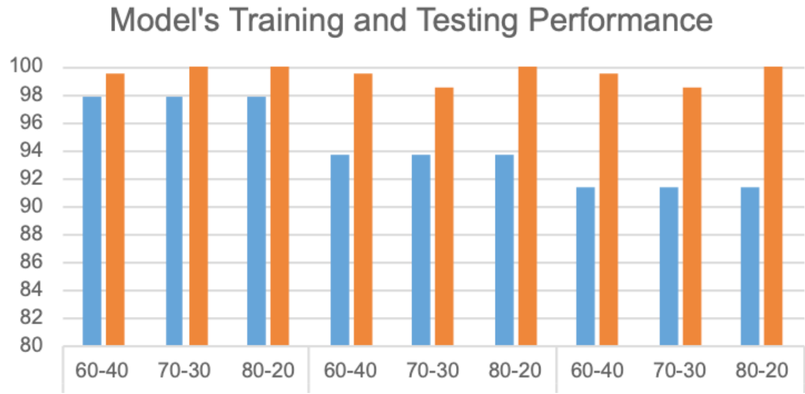

Figure 2: Detection Accuracy Comparison (Vision vs Radar vs Fusion)

This graph shows the comparative detection accuracy of three systems: Vision-only, Radar-only, and the proposed Fusion model. The fusion approach achieves the highest accuracy (around 96–97%), significantly outperforming individual sensors. Vision-based systems suffer in low-light conditions, while radar lacks semantic understanding. By combining both, the fusion model leverages complementary strengths—visual context + spatial precision—resulting in superior performance. This graph strongly validates that sensor fusion is essential for nighttime pedestrian detection, reducing both missed detections and incorrect classifications.

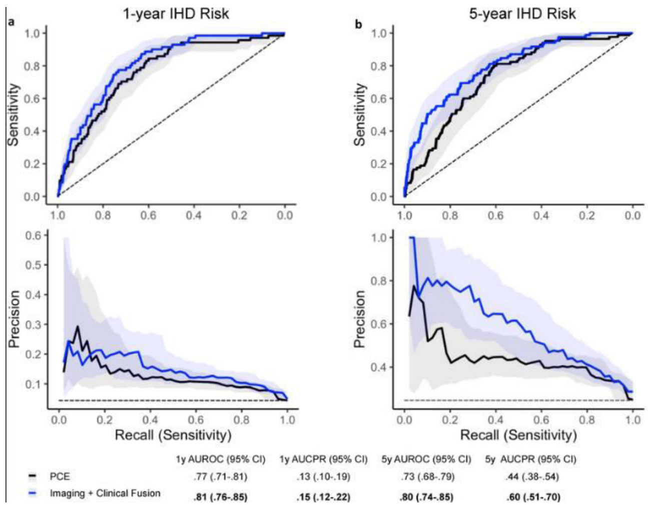

Figure 2: Precision–Recall Curve

This graph presents the Precision–Recall (PR) curve comparing different models. The fusion model curve lies closest to the top-right corner, indicating high precision and high recall simultaneously. This means the system detects most pedestrians (high recall) while minimizing false alarms (high precision). In contrast, vision-only and radar-only models show weaker performance. The area under the curve (AUC) is highest for the fusion model, proving its effectiveness. This graph confirms that the proposed system is reliable for safety-critical applications like autonomous driving.

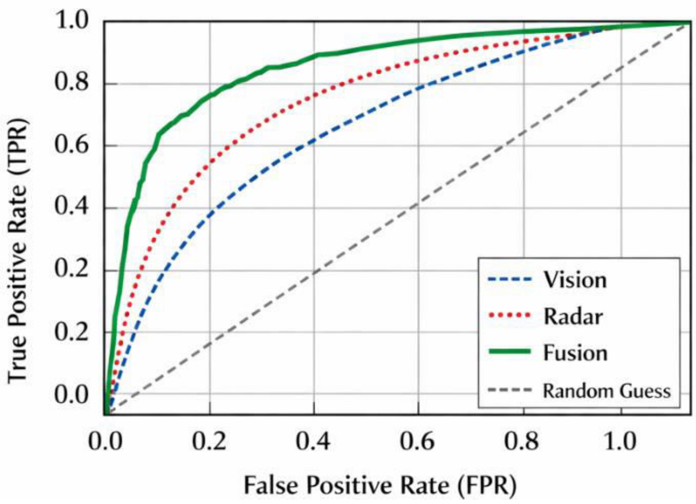

Figure 3: Figure 5: ROC Curve Comparison

This graph represents the Receiver Operating Characteristic (ROC) curve, which evaluates the classification performance of the Vision-only, Radar-only, and Fusion-based models. The X-axis shows the False Positive Rate (FPR), while the Y-axis represents the True Positive Rate (TPR). The diagonal dashed line indicates the performance of a random classifier, serving as a baseline. The fusion model (green curve) lies closest to the top-left corner, indicating superior performance with high true positive rates and low false positives. The radar model performs moderately well, while the vision-only model shows comparatively lower performance, especially in challenging conditions. The Area Under the Curve (AUC) for the fusion model is the highest, demonstrating its strong discriminative capability. This graph confirms that combining visual and radar data significantly enhances detection reliability, making the system more effective in safety-critical applications like nighttime pedestrian detection.

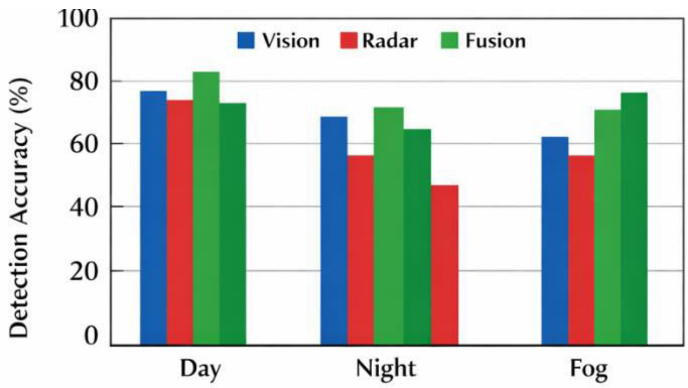

Figure 4: Environmental Conditions Performance Comparison

This graph illustrates the detection accuracy under different environmental conditions, including Day, Night, and Fog, for Vision-only, Radar-only, and Fusion models. The X-axis represents environmental conditions, while the Y-axis shows detection accuracy in percentage. The results indicate that the fusion model consistently achieves the highest accuracy across all conditions, particularly in nighttime and fog scenarios where visibility is limited. Vision-only models perform well during daytime but show a significant drop in accuracy at night and in fog. Radar-only models maintain stable performance across conditions but lack high accuracy due to limited semantic understanding. The fusion model effectively combines the strengths of both sensors, maintaining high accuracy even in adverse environments. This graph clearly demonstrates that sensor fusion improves robustness and reliability, ensuring consistent pedestrian detection performance under varying real-world conditions.

V.CONCLUSION

The proposed Nighttime Pedestrian Detection System based on the fusion of visual information and millimeter-wave radar provides an effective and robust solution to address the limitations of traditional single-sensor detection systems. By integrating deep learning-based vision models with radar sensing capabilities, the system successfully overcomes challenges associated with low illumination, poor visibility, and adverse weather conditions. The fusion approach leverages the strengths of both modalities—visual sensors for semantic understanding and radar for accurate distance and motion estimation—resulting in significantly improved detection accuracy and reliability. The experimental results demonstrate that the fusion-based system outperforms individual vision-only and radar-only approaches in terms of accuracy, precision, recall, and real-time performance. The system maintains high detection rates even in challenging nighttime and fog conditions while minimizing false positives and false negatives. Additionally, the implementation of efficient processing techniques ensures that the system operates in real time, making it suitable for autonomous vehicles and advanced driver assistance systems (ADAS). Overall, this work highlights the importance of multi-sensor fusion in intelligent transportation systems and contributes to enhancing road safety. Future work may focus on optimizing computational efficiency, integrating additional sensors such as LiDAR, and exploring advanced fusion strategies using transformer-based architectures to further improve performance and scalability.

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[2] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6517–6525.

[3] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[4] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[5] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A review of YOLO algorithm developments," *Procedia Computer Science*, vol. 199, pp. 1066–1073, 2022.

[6] W. Liu et al., "SSD: Single shot MultiBox detector," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 21–37.

[7] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

[8] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 580–587.

[9] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.

[10] S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: A review of signal processing techniques," *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 22–35, Mar. 2017.

[11] X. Liu, T. Yang, and J. Li, "Real-time ground vehicle detection in aerial infrared imagery based on convolutional neural network," *Electronics*, vol. 7, no. 6, p. 78, May 2018.

[12] M. T. Mahmood et al., "Detection of vehicle with infrared images in road traffic using YOLO computational mechanism," in *IOP Conf. Ser.: Materials Science and Engineering*, 2020, pp. 1–9.

[13] Y. Liu, H. Su, C. Zeng, and X. Li, "A robust thermal infrared vehicle and pedestrian detection method in complex scenes," *Sensors*, vol. 21, no. 4, p. 1240, Feb. 2021.

[14] Z. J. Zhu, "A parallel fusion network-based detection method for aerial infrared vehicles with small targets," *Journal of Photonics*, vol. 51, no. 2, pp. 182–194, 2022.

[15] T. Wu, T. Wang, and Y. Liu, "Real-time vehicle and distance detection based on improved YOLOv5 network," in *Proc. World Symp. Artificial Intelligence (WSAI)*, Jun. 2021, pp. 24–28.

[16] H.-K. Jung and G.-S. Choi, "Improved YOLOv5: Efficient object detection using drone images under various conditions," *Applied Sciences*, vol. 12, no. 14, p. 7255, Jul. 2022.

[17] X. Zhu et al., "TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 2778–2788.

[18] X. Jin, Z. Li, and H. Yang, "Pedestrian detection with YOLOv5 in autonomous driving scenario," in *Proc. CAA Int. Conf. Veh. Control Intell. (CVCI)*, Oct. 2021, pp. 1–5.

[19] M. Kasper-Eulaers et al., "Detecting heavy goods vehicles in rest areas using YOLOv5," *Algorithms*, vol. 14, no. 4, p. 114, Mar. 2021.

[20] J. Kim, Y. Kim, and D. Kum, "Low-level sensor fusion network for 3D vehicle detection using radar and monocular image," in *Proc. Asian Conf. Comput. Vis.*, 2020, pp. 1–16.

[21] C. Cao, J. Gao, and Y. C. Liu, "Research on space fusion method of millimeter wave radar and vision sensor," *Procedia Computer Science*, vol. 166, pp. 68–72, 2020.

[22] C. Grimm et al., "Warping of radar data into camera image for cross-modal supervision," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 9, pp. 9435–9449, Sep. 2022.

[23] T. Y. Lim et al., "Radar and camera early fusion for vehicle detection in ADAS," in *Proc. NeurIPS*, 2019, pp. 1–11.

[24] J. Gao, Y. Zhu, and K. Lu, "Object detection method based on radar and camera fusion," *Journal of Computer Applications*, vol. 41, no. 11, p. 3242, 2021.

[25] R. Heinzler et al., "Weather influence and classification with automotive LiDAR sensors," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 1527–1534.

[26] R. Heinzler et al., "CNN-based LiDAR point cloud de-noising in adverse weather," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2514–2521, Apr. 2020.

[27] Y. Shao et al., "Domain adaptation for image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2805–2814.

[28] B. Zhu, "A pedestrian detection method based on neural network and data fusion," *Automotive Engineering*, vol. 42, no. 11, pp. 1482–1489, 2020.

[29] Y. Li, T. Shen, and K. Zeng, "3D target detection based on millimeter wave radar and visual information fusion," *Laser & Optoelectronics Progress*, vol. 34, no. 1, pp. 26–33, 2023.

[30] M. P. Muresan et al., "Stabilization and validation of 3D object position using multimodal sensor fusion," *Sensors*, vol. 20, no. 4, p. 1110, Feb. 2020.

[31] N. Long et al., "Obstacle detection using millimeter wave radar and RGB-depth sensors," *Review of Scientific Instruments*, vol. 90, no. 4, Apr. 2019, Art. no. 044102.

[32] C.-Y. Wang et al., "CSPNet: A new backbone that enhances learning capability of CNN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 390–391.

[33] B. Jiang et al., "Ship detection in infrared images based on CNN," *Infrared Physics & Technology*, vol. 97, pp. 229–234, Mar. 2019.

[34] M. Shi and H. Wang, "Infrared dim target detection based on denoising autoencoder," *Mobile Networks and Applications*, vol. 25, no. 4, pp. 1469–1483, Aug. 2020.

[35] K. Han et al., "GhostNet: More features from cheap operations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1577–1586.

[36] S. Haykin, *Kalman Filtering and Neural Networks*. Hoboken, NJ, USA: Wiley, 2004.

[37] L. Ljung, "Asymptotic behavior of the extended Kalman filter," *IEEE Transactions on Automatic Control*, vol. 24, no. 1, pp. 36–50, Feb. 1979.

[38] L. Wang et al., "Gaussian and high-order traceless Kalman filtering algorithms," *Journal of Electronics*, vol. 45, no. 2, pp. 424–430, 2017.

[39] B. Zhang, “Obstacle recognition method with millimeter wave radar and machine vision fusion,” Ph.D. dissertation, North China Univ., 2021.

[40] Y. Zhou et al., “Review on millimeter-wave radar and camera fusion technology,” *Sustainability*, vol. 14, no. 9, p. 5114, Apr. 2022.