Magna Scientia Advanced Research and Reviews
eISSN: 2582-9394
Cross Ref DOI: 10.30574/msarr
Journal homepage: https://magnascientiapub.com/journals/msarr/

(Review Article)

Check for updates

# A review of multisensor data fusion techniques for reliable autonomous perception

Derrick Appiah Osei [1] and Zakaria Yakin [2,*]

[1] *Drexel University, U.S.A.*
[2] *Kwame Nkrumah University of Science and Technology, Ghana.*

Magna Scientia Advanced Research and Reviews, 2026, 17(01), 125-130

Publication history: Received on 13 April 2026; revised on 20 May 2026; accepted on 22 May 2026

Article DOI: https://doi.org/10.30574/msarr.2026.17.1.0086

## Abstract

Autonomous systems such as self-driving vehicles, drones, and intelligent robots rely on accurate perception to interact safely with dynamic environments. However, single-sensor systems are prone to noise, occlusion, and environmental variability, limiting their reliability. Multisensor Data Fusion (MSDF) has therefore emerged as a crucial approach for improving perception by integrating complementary information from multiple sensors, including cameras, LiDAR, radar, GPS, and inertial measurement units (IMUs). This paper reviews current MSDF techniques and their role in achieving reliable, robust, and real-time autonomous perception. The review categorizes fusion approaches into three major groups: classical probabilistic models, knowledge-based reasoning techniques, and data-driven deep learning methods. Classical models such as Kalman Filters, Extended Kalman Filters, Particle Filters, and Bayesian networks provide strong statistical foundations for uncertainty estimation and sensor integration. Knowledge-based systems, including Dempster–Shafer theory and fuzzy logic, effectively manage imprecise or conflicting sensory information. More recently, deep learning-based fusion techniques, leveraging Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and attention-based models, have enabled end-to-end learning from multimodal inputs, achieving higher perception accuracy and adaptability in complex environments. A comparative analysis reveals that classical fusion methods offer transparency and low computational cost but are limited in scalability and nonlinearity handling. Deep learning approaches, while powerful, face challenges related to explainability, computational demand, and data dependency. Finally, the study highlights emerging trends shaping the future of autonomous perception, including edge AI and distributed fusion, explainable AI (XAI) for decision transparency, and the integration of 5G/6G communication for cooperative sensing.

**Keywords:** Multisensor Data Fusion; Autonomous Perception; Deep Learning; Uncertainty Quantification; Reliability; Robotics; Intelligent Systems

## 1. Introduction

Reliable perception is fundamental to the safe operation of autonomous systems, including self-driving vehicles, unmanned aerial vehicles (UAVs), and intelligent robots. Every autonomous platform depends on accurate and timely environmental information for navigation, object avoidance, and decision-making. However, perception based on a single sensor such as a camera or LiDAR remains susceptible to occlusion, adverse weather, sensor noise, and limited field of view, often resulting in incomplete or inaccurate situational awareness [1]. These limitations can significantly compromise system reliability and safety, particularly in complex and dynamic environments.

To mitigate these weaknesses, multisensor data fusion (MSDF) has become a cornerstone of modern autonomous perception. By integrating complementary information from sensors such as cameras, LiDAR, radar, GPS, and inertial measurement units (IMUs), MSDF enables redundancy, compensates for individual sensor deficiencies, and enhances perception robustness [2]. For instance, cameras provide rich semantic information, LiDAR contributes accurate depth

* Corresponding author: Zakaria Yakin

geometry, radar offers strong performance under poor visibility, and IMUs supply motion continuity. When combined, these modalities support essential perception tasks object detection and classification, 3-D localization, mapping, and semantic scene understanding while providing resilience against partial sensor failure [3]. MSDF techniques span a broad methodological spectrum. At the classical end, probabilistic and statistical filters such as the Kalman Filter (KF), Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), and Particle Filter have long been used for sensor fusion and state estimation due to their rigorous uncertainty modeling and real-time efficiency [4]. These approaches are effective when the dynamic model and noise distributions are known or can be well approximated. Complementary methods including Dempster–Shafer evidential reasoning and fuzzy logic systems are useful for handling uncertain or conflicting sensor information, particularly when measurements disagree due to occlusion or interference [5].

In recent years, deep learning-based fusion has transformed autonomous perception by learning complex nonlinear relationships directly from data. Convolutional neural networks (CNNs), recurrent networks (RNNs), and transformer architectures enable powerful feature-level and decision-level fusion. These data-driven models combine raw or intermediate representations from heterogeneous sensors to improve spatial and semantic understanding [3]. For example, CNN-LiDAR fusion networks have demonstrated substantial improvements in object detection accuracy and localization precision compared with traditional rule-based methods [6]. However, these models require large labeled datasets, impose heavy computational demands, and pose challenges for interpretability and safety certification [1]. Reliability and uncertainty management remain critical to perception performance. Research has emphasized probabilistic deep learning, Bayesian neural networks, and Monte-Carlo dropout techniques to quantify prediction uncertainty and enable safer decision-making [6]. Additionally, fault-tolerant fusion architectures have been developed to maintain operational safety during sensor degradation or failure [4]. These advances align with the broader movement toward certifiable, trustworthy AI in safety-critical domains. From a systems perspective, practical deployment of MSDF must address latency, computational resources, and communication constraints. The growing adoption of edge AI and distributed fusion frameworks enables real-time inference by performing computation near the sensor source, reducing bandwidth requirements and improving responsiveness [1]. At the same time, the emergence of explainable AI (XAI) is enhancing transparency and user trust by clarifying how multimodal fusion models derive decisions [7]. These trends illustrate the convergence of perception, connectivity, and intelligence in the next generation of autonomous systems.

## 2. Fundamentals of Multisensor Data Fusion

Multisensor data fusion in autonomous perception involves the integration of heterogeneous measurements such as camera imagery, LiDAR point clouds, radar Doppler returns, and IMU/GNSS signals to build a more complete, robust and actionable environmental model than any single modality alone [8]. Conceptually, fusion is framed by three levels of abstraction: signal (or measurement) level, feature level, and decision level. At the signal level, raw or lightly pre-processed sensor streams are directly fused to improve signal quality and state estimation; at the feature level, intermediate representations (for example keypoints, voxels or bird's-eye-view features) are aligned and combined to exploit complementary modality cues; and at the decision level, independent hypotheses (such as object tracks or semantic decisions) are merged to form a final inference or decision [9; 10]. Samadzadegan et al. further note that within remote sensing fusions, feature-level approaches dominate across over 950 surveyed studies, outnumbering pixel/measurement and decision-level fusions [10].

In terms of architecture, two broad configurations are contrasted: centralized and decentralized fusion. In a centralized architecture, all sensor measurements (or derived features) are transmitted to and processed at a common fusion hub, maximizing information use but incurring higher bandwidth, compute and communication costs; in a decentralized architecture, edge nodes (e.g., sensors or local ECUs) locally preprocess or infer state estimates, then share summaries or tracks to a fusion center, reducing communication burden and enabling graceful degradation under partial failures [8]. More advanced frameworks adopt hierarchical or hybrid designs: for example, fast low-latency tasks (ego-motion, obstacle detection) may be processed at edge nodes, while global scene reasoning is reserved for higher tiers, and hybrid fusion may combine both early (raw/feature) and late (decision) fusion within one pipeline [10, 11].

Key performance criteria for fusion systems in autonomous perception extend beyond accuracy to include reliability, computational efficiency and real-time adaptability. Accuracy pertains to standard metrics such as detection/segmentation AP or localization RMSE; reliability refers to system robustness, calibration integrity, uncertainty calibration and fault tolerance; computational efficiency addresses throughput, latency, memory and energy budgets; and real-time adaptability reflects the system's ability to meet operational deadlines [8; 12]. For practical deployment in moving platforms, it has been highlighted that representation-efficient fusion and edge-placement of compute are critical for meeting cycle-time constraints while preserving robustness [12]. Despite these advances, autonomous environments which include mobile, dynamic, sensor-rich platforms persistent fusion challenges. Sensor uncertainty (such as lighting/ weather domain shift, calibration drift) can propagate bias into fused outputs; synchronization across asynchronous modalities remains non-trivial; data redundancy must be managed such that complementary cues are used without overwhelming resources; and communication constraints (limited bandwidth, intermittent links) force trade-offs between raw/feature/summary transmission, compression and prioritization [11; 8]. Moreover, architectural studies show that decentralized/hierarchical fusion with edge offloading improves resilience under packet loss or load spikes but still demands robust calibration and temporal alignment. Continued research is needed in adaptive fusion algorithms that can down-weight degraded sensors on the fly, monitor calibration drift, and sustain real-time behaviour in constrained networks [10, 12].

## 3. Classical Multisensor Data Fusion Techniques

Classical multisensor data fusion methods form the analytical foundation for integrating diverse sensor observations into a consistent representation of the environment. Probabilistic techniques are among the most established, with the Kalman Filter (KF) providing optimal recursive estimation for linear Gaussian systems, and its variants the Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) addressing nonlinear and higher-order dynamics. These methods remain central to vehicle localization and object tracking in autonomous systems. When the underlying distributions are non-Gaussian or highly nonlinear, Particle Filters (PFs) approximate posterior probabilities using weighted samples, offering robustness in complex dynamic environments. All of these methods are rooted in Bayesian inference, which enables continuous belief updating as new sensor data are received, ensuring statistically coherent integration of uncertain information [8].

Beyond probabilistic models, evidential reasoning and fuzzy logic frameworks address uncertainty that cannot be readily modeled probabilistically. The Dempster–Shafer Theory (DST) allows belief mass assignment to sets of hypotheses, effectively handling conflicting or incomplete sensor data. In contrast, fuzzy logic systems use membership functions and rule-based reasoning to interpret imprecise sensor readings, such as partial visibility or vague confidence levels, making them suitable for real-world perception tasks with imperfect inputs [9]. These approaches complement probabilistic methods by providing greater flexibility when data are ambiguous or incomplete.

Statistical and optimization-based techniques also play a key role in classical data fusion. Least-squares estimation minimizes measurement discrepancies, providing efficient solutions for sensor alignment and calibration, while the Covariance Intersection (CI) method offers conservative fusion when cross-correlation between sensors is unknown. Markov Random Fields (MRFs) extend statistical modeling to spatially dependent data, making them effective in image and map fusion. Comparative analyses indicate that probabilistic methods generally excel in dynamic estimation, evidential and fuzzy systems in qualitative reasoning, and optimization-based techniques in computational efficiency, especially under real-time constraints [12]. Together, these classical techniques underpin modern hybrid and learning-based fusion frameworks that drive reliable perception in autonomous systems.

## 4. Machine Learning and Deep Learning-Based Fusion

The rise of machine learning (ML) and deep learning (DL) has transformed multisensor data fusion, enabling systems to learn nonlinear, high-dimensional relationships directly from data. Unlike classical probabilistic models, which rely on handcrafted features and explicit noise models, ML- and DL-based approaches learn both features and fusion rules jointly through optimization. Feature-level fusion is often achieved using Convolutional Neural Networks (CNNs) that combine visual texture from cameras with spatial geometry from LiDAR to improve perception in autonomous driving [11]. CNN-based fusion architectures, such as BEVFusion and PointPainting, project sensor modalities into a shared latent space before joint reasoning, yielding significant gains in detection accuracy under challenging lighting or weather. For dynamic environments, Recurrent Neural Networks (RNNs) and Transformers extend this capability to the temporal domain, modeling sequential dependencies and motion cues for tracking and trajectory prediction [13]. These models leverage self-attention to align temporal patterns across asynchronous sensor streams, improving robustness to occlusions and sensor delays.

At the decision level, ensemble learning methods integrate outputs from multiple base learners to enhance prediction reliability. Techniques such as Random Forests, Gradient Boosting Machines, and hybrid voting schemes aggregate independent classifier results from different sensors or subsystems. These models are valued for their interpretability and low training cost compared to deep neural networks while offering robustness to noise and partial sensor failures [12]. Recent advances in end-to-end multimodal learning have enabled networks to process raw sensor data jointly without explicit feature engineering. Multimodal autoencoders learn shared latent representations across modalities, while Generative Adversarial Networks (GANs) simulate missing modalities to enhance robustness under partial data.

Multimodal attention networks further improve interpretability by dynamically weighting each modality according to contextual relevance [14]. These architectures allow perception models to adaptively emphasize reliable sensors while suppressing degraded ones, a key requirement for safety-critical autonomous systems.

**Table 1** Comparative Analysis of Machine Learning and Deep Learning Fusion Approaches (Prusty et al., 2023 and Chen et al., 2024)

| Fusion Approach | Key Strengths | Limitations | Performance under Sensor Failure | Interpretability | Generalization |
|---|---|---|---|---|---|
| CNN-based Feature Fusion | Strong spatial perception; integrates texture + depth | Requires high-quality calibration; computationally intensive | Moderate fails gracefully if LiDAR or camera missing | Medium (visualizable feature maps) | High for similar domains |
| RNNs / Transformers (Temporal Fusion) | Captures motion and sequence context | Sensitive to synchronization errors | High attention helps re-weight valid sensors | Low–Medium | High if pretrained on diverse data |
| Ensemble Learning (Decision-Level) | Robust, lightweight, interpretable | Limited nonlinear feature learning | High redundant classifiers mitigate loss | High | Moderate |
| End-to-End Deep Fusion (Autoencoders / GANs / Attention) | Learns directly from raw multimodal data; adaptive | Data-hungry; less explainable | Very high attention and GAN imputation handle missing inputs | Low | High with large datasets |

## 5. Reliability and Uncertainty Management in Data Fusion

Ensuring reliability and safety in multisensor data fusion is critical for the deployment of autonomous systems in dynamic, unpredictable environments. As perception pipelines increasingly depend on complex sensor arrays cameras, LiDAR, radar, GPS, and inertial units the ability to handle degradation, quantify uncertainty, and maintain operational safety has become a central research focus [12]. Sensor reliability models are designed to evaluate the health and trustworthiness of individual sensors within the fusion framework. They account for degradation over time, temporary occlusions, and permanent failures, assigning reliability scores that inform the weighting of sensor contributions. Probabilistic reliability modeling and confidence-based fusion enable systems to down-weight unreliable sources in real time, improving resilience to partial sensor dropout [13]. In autonomous driving, for example, reliability maps generated from LiDAR reflectivity or camera signal-to-noise ratios help adaptively adjust fusion parameters during poor weather or low-visibility conditions.

A closely related issue is uncertainty quantification, which aims to measure how confident a fusion system is about its predictions. Classical probabilistic graphical models express epistemic and aleatoric uncertainties by modeling variable dependencies explicitly. Modern deep learning-based systems achieve similar objectives through Monte Carlo dropout or Bayesian deep learning, where uncertainty is estimated by sampling model weights during inference. These approaches allow perception modules to convey calibrated confidence levels to downstream decision-making and control systems essential for safe trajectory planning in autonomous vehicles [14].

Reliability also depends on fault tolerance and redundancy mechanisms, which ensure continuous operation under sensor or subsystem failure. Methods such as Covariance Intersection, ensemble fusion, and multimodal attention networks can re-route decision flows when sensors malfunction or provide inconsistent readings. Redundant sensing modalities (e.g., radar + LiDAR) are often fused through adaptive weighting or Bayesian evidence frameworks, maintaining accurate state estimation even when one modality fails [11]. Finally, these concepts underpin safety-critical applications such as autonomous driving, aerial navigation, and robotic manipulation. In self-driving vehicles, multimodal fusion frameworks combine LiDAR geometry with visual semantics to detect obstacles under challenging weather, while UAV navigation systems integrate inertial, GPS, and visual odometry data to ensure stable flight during signal loss [8]. Real-world validations consistently show that reliable fusion with quantified uncertainty enhances operational safety, interpretability, and trustworthiness cornerstones of dependable autonomous systems.

## 6. Research Gaps and Future Directions

Despite significant advancements in multisensor fusion for autonomous perception, several research gaps remain that constrain real-world reliability, scalability, and interpretability. One emerging direction is the integration of physics-informed learning and probabilistic reasoning. While deep neural networks excel at pattern recognition, they often disregard the physical constraints and causal relationships underlying sensor data. Embedding physics-based priors into neural architectures through hybrid physics-informed neural networks (PINNs) can enhance model interpretability, reduce data dependence, and improve generalization across domains. Another challenge lies in the standardization of multisensor datasets and benchmarks. Current datasets vary widely in sensor configurations, annotation quality, and environmental diversity, making fair comparison and reproducibility difficult. Establishing unified benchmarks for multimodal fusion covering diverse weather, lighting, and terrain conditions would accelerate progress and support consistent model evaluation.

The need for real-time fusion under resource constraints represents a third critical research frontier. Autonomous systems deployed in embedded or edge environments must balance computational efficiency with accuracy and robustness. Efficient architectures, such as sparse attention models, quantized neural networks, and adaptive fusion pipelines, are being explored to achieve real-time inference without sacrificing safety. In addition to technical concerns, ethical and privacy considerations are gaining prominence. The integration of multimodal sensors, including cameras and LiDAR, raises concerns about surveillance, data ownership, and unintended bias in perception algorithms. Responsible fusion research must incorporate privacy-preserving mechanisms such as federated learning or differential privacy to ensure equitable and ethical deployment.

Finally, there is an urgent call for reliable certification pathways for autonomous systems. Unlike traditional software validation, the stochastic nature of learning-based fusion systems challenges existing regulatory frameworks. Developing standardized validation protocols that incorporate reliability metrics, uncertainty bounds, and fault tolerance testing is essential for public trust and large-scale deployment.

## 7. Conclusion

This review highlights that multisensor data fusion remains a cornerstone of reliable autonomous perception. Classical probabilistic models such as Kalman filters and Bayesian inference laid the groundwork for robust state estimation, while modern deep learning methods have introduced data-driven feature extraction, temporal reasoning, and adaptive decision-making capabilities. Hybrid approaches that combine these paradigms now dominate the frontier, offering improved accuracy, resilience, and contextual awareness. Nevertheless, ensuring trustworthy autonomy extends beyond accuracy alone. The integration of uncertainty quantification, redundancy mechanisms, and real-time reliability assessment is vital to sustain performance under sensor degradation or failure. Future systems will need to fuse insights from AI, robotics, control theory, and human–machine interaction, forming interdisciplinary frameworks that can reason under uncertainty while adhering to ethical and safety standards. As autonomous vehicles, drones, and robotic systems proliferate, advances in multisensor fusion will continue to define the next generation of intelligent, context-aware, and resilient machines capable of operating safely in complex environments.

## Compliance with ethical standards

### *Disclosure of conflict of interest*

No conflict of interest to be disclosed.

## References

[1] Yeong, D. J., Velasco-Hernandez, G., Barry, J., and Walsh, J. (2025). Sensor and sensor fusion technology for autonomous vehicles: A review. Sensors, 25(1), 198. https://doi.org/10.3390/s25010198

[2] Du, J., Zhao, L., and Sun, Y. (2024). A hybrid LiDAR–camera fusion approach for enhanced autonomous perception under low-visibility conditions. IEEE Transactions on Intelligent Transportation Systems, 25(3), 2841–2856. https://doi.org/10.1109/TITS.2024.3350021

[3] Tan, Z., Wang, C., and Li, Q. (2024). Cross-modal fusion networks for robust 3D object detection in autonomous driving. IEEE Robotics and Automation Letters, 9(1), 1021–1028. https://doi.org/10.1109/LRA.2024.3342069

[4] Liu, J., Zhou, J., and Han, M. (2023). A review of classical Kalman-based and particle filter algorithms for autonomous navigation. Sensors, 23(4), 1672. https://doi.org/10.3390/s23041672

[5] Gupta, P., and Sharma, R. (2022). Evidential reasoning-based multisensor fusion for robust environmental perception. Information Fusion, 84, 110–125. https://doi.org/10.1016/j.inffus.2022.01.013

[6] Fan, H., Chen, Y., and Wang, J. (2025). Bayesian deep learning for uncertainty-aware perception in autonomous vehicles. IEEE Transactions on Neural Networks and Learning Systems, 36(2), 1508–1521. https://doi.org/10.1109/TNNLS.2025.3375012

[7] Zhang, Y., Li, C., and Xu, D. (2023). Explainable AI for sensor fusion in autonomous systems. Pattern Recognition Letters, 174, 56–64. https://doi.org/10.1016/j.patrec.2023.05.012

[8] Duro, N. (2024). Sensor data fusion analysis for broad applications. Sensors, 24(12), 3725. https://doi.org/10.3390/s24123725

[9] Li, B., Xian, Y., Zhang, D., Su, J., Hu, X., and Guo, W. (2021). Multi-sensor image fusion: A survey of the state of the art. Journal of Computer and Communications, 9(6), 73–108. https://doi.org/10.4236/jcc.2021.96005

[10] Samadzadegan, F., Toosi, A., and Dadrass Javan, F. (2025). A critical review on multi-sensor and multi-platform remote sensing data fusion approaches: Current status and prospects. International Journal of Remote Sensing, 46(3), 1327–1402. https://doi.org/10.1080/01431161.2024.2429784

[11] Huang, K., Shi, B., Li, X., Li, Y., and Huang, S. (2022). Multi-modal sensor fusion for auto-driving perception: A survey. IEEE Access, 10, 118321–118345. https://doi.org/10.1109/ACCESS.2022.3203341

[12] Ma, X., Zhou, P., He, X., and Zhang, S. (2025). A comprehensive review of multi-source data fusion processing methods. Preprints. https://doi.org/10.20944/preprints202503.1108.v1

[13] Qian, R., Li, H., and Zhao, Y. (2025). Multimodal sensor fusion for autonomous driving: Recent trends and future directions. Sensors, 25(4), 1415. https://doi.org/10.3390/s25041415

[14] Zhang, L., and He, J. (2024). Attention-based deep fusion networks for multimodal perception. IEEE Sensors Journal, 24(8), 9871–9883. https://doi.org/10.1109/JSEN.2024.3367829