OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 03 - Issue 03 (2026)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

SAGE-IOT: A LIGHTWEIGHT LLM-INSPIRED FRAMEWORK FOR MULTIMODAL FUSION AND BAYESIAN DECISION-MAKING IN IOT

***Qifan Shen*[1], *Qirong Shen*[2], *Lei Sun*[3], *Sihan Shen*[4], **Yutong Ye*[5], *Muhammad Wahab Hanif*[6]**
[1, 2, 3, 4, 5, 6]*Foretek Smart Technology (Zhejiang) Co., Ltd. Shaoxing, Zhejiang Province, 312030, China*

[*]***Corresponding Author:*** (dayle.ye@hotmail.com)
DOI: (*https://doi.org/10.71146/kjmr1007*)

Article Info

Abstract

This paper proposes SAGE-IoT, a lightweight, large-language-model-inspired multimodal decision-making framework for addressing two persistent challenges in Internet of Things (IoT) scenarios: the difficulty of jointly modeling logs, images, and sensor time series, and the asymmetric costs of decision errors. Drawing on the ideas behind Parameter-Efficient Fine-Tuning (PEFT), Knowledge-Augmented Reasoning (KAR), and hierarchical prompting, the framework implements a "multimodal encoding → gated fusion → knowledge-enhanced reasoning → Bayesian action selection → closed-loop feedback" pipeline that strengthens edge decision-making without deploying a complete Large Language Model (LLM). Experiments across three simulated scenarios, smart home, industrial fault warning, and smart transportation, achieve accuracy of 94.6%, 88.4%, and 91.9%, respectively. Ablation results show that KAR and PEFT contribute gains of approximately 3.8 and 1.7 percentage points, respectively. These results indicate that lightweight LLM-inspired mechanisms offer a feasible way to enhance traditional multimodal models, and that "calibrate before deciding" should be a basic design principle for intelligent IoT decision-making systems.

Keywords: *Edge Intelligence; Multimodal Data Fusion; Parameter-Efficient Fine Tuning; Risk Aware Decision Making*

1. Introduction

Internet of Things (IoT) systems must simultaneously process text logs, visual information, and sensor time series. However, these three data types differ markedly in structure, frequency, and semantic granularity, which makes traditional rule-based systems difficult to scale and leaves purely deep-learning models without an interpretable basis for decision-making. In recent years, Large Language Models (LLMs) have shown strong performance in complex reasoning, prompt organization, and knowledge invocation, offering new ideas for unifying perception and decision-making[1], [2], [3], [4]. At the same time, techniques such as parameter-efficient fine-tuning, quantized adaptation, and retrieval augmentation have substantially lowered the barrier to deployment, opening a feasible path toward edge intelligence[5], [6], [7]. Recent work has also begun to explore how LLMs can be connected to the physical world, multimodal understanding, and IoT tasks[8], [9], [10], [11], [12], [13]. However, systematic research on closed-loop IoT decision-making remains limited, and a unified framework is still lacking, particularly for joint modeling of heterogeneous modalities, control of edge-deployment cost, and cost-sensitive decision-making in high-risk scenarios.

IoT decision-making is subject to three constraints: sensory data arrive asynchronously, edge computing power is limited, and the costs of different types of errors are markedly asymmetric. A practical system must therefore achieve accurate recognition, interpretable decisions, deployability, and controllable action costs all at once. Based on this observation, rather than deploying a complete LLM, this paper distills the key ideas behind Parameter-Efficient Fine-Tuning (PEFT), Knowledge-Augmented Reasoning (KAR), and hierarchical prompts to build SAGE-IoT, a framework that can run within a conventional machine-learning compute budget.

The contributions of this paper are threefold: (1) a closed-loop multimodal decision-making architecture designed for IoT; (2) an integration of PEFT, KAR, and gated fusion that improves decision performance in few-shot and heterogeneous settings; and (3) the introduction of Bayesian risk decision-making, together with experiments that demonstrate the necessity of probability calibration[14], [15].

2. SAGE-IoT System Design

2.1. Overall Architecture

The SAGE-IoT framework consists of five layers, and its overall architecture is shown in Figure 1.

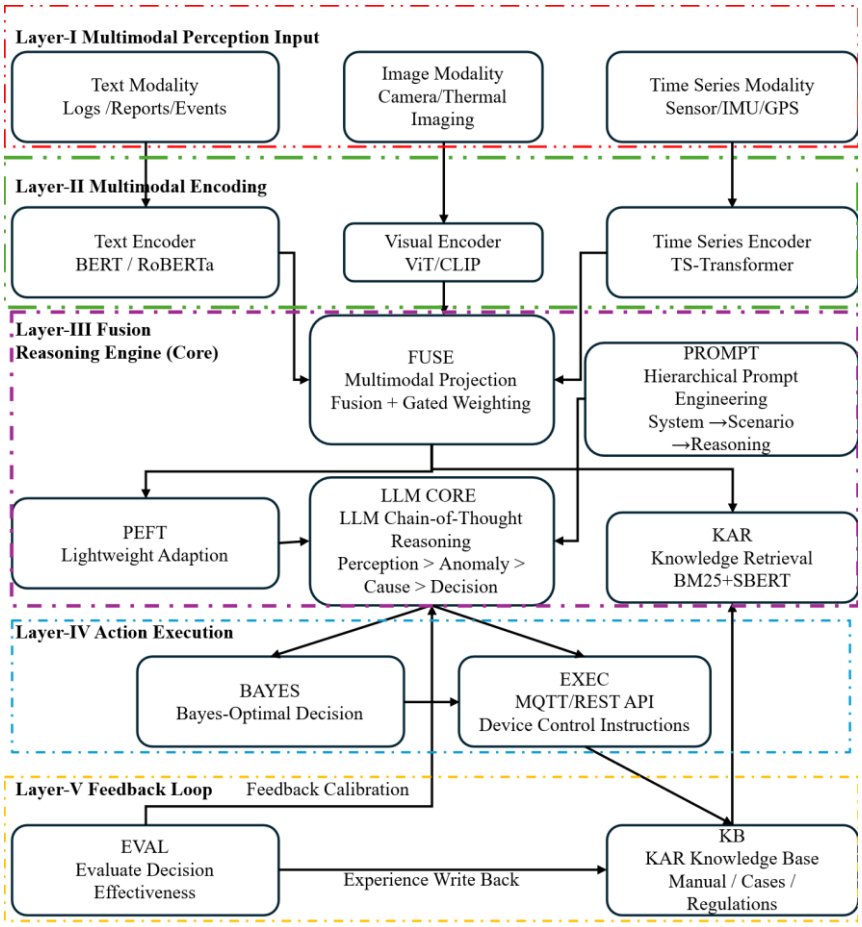

Figure 1. SAGE-IoT framework

SAGE-IoT is composed of five layers: perception, encoding, fusion reasoning, action execution, and feedback. The perception layer receives text, image, and sensor data; the encoding layer extracts features separately for each of the three modalities; the fusion reasoning layer performs projection and gated weighting, and combines PEFT adaptation with KAR retrieval to produce scenario-level decisions shown in Equation 1 to Equation 4; the action-execution layer generates control instructions according to a risk-minimization principle; and the feedback layer writes execution results back to the knowledge base, forming a continuous optimization loop[16], [17], [18], [19], [20]. This architecture brings knowledge retrieval, semantic reasoning, and cost constraints into a single pipeline, making it well suited to complex IoT scenarios.

2.2. Key Methods and Algorithms of Each Layer

2.2.1. Perception layer (acquisition and preprocessing)

The input consists of three types of raw signals: text, image, and sensor data. The perception layer is responsible for transcription, denoising, standardization, and temporal alignment, preparing the data for consumption by the encoders.

2.2.2. Encoding layer (modality feature extraction)

Text, image, and time-series encoders are used to extract local semantic and temporal features, respectively, all uniformly mapped to dimension d_h. The resulting outputs are denoted as the text-modality encoding vector z_t, the image-modality encoding vector z_i, and the sensor time-series encoding vector z_s.

2.2.3. Fusion reasoning layer (Projection → Gating → PEFT → KAR)

Let the multimodal projection-fusion vector of the three modality encodings, h_{fused} be defined as

$$h_{\text{fused}} = W_{\text{proj}} [z_t \| z_i \| z_s] + b_{\text{proj}} \quad (1)$$

here denotes concatenation, $W_{\text{proj}} \in R^{d_h \times 3d_h}$ is a learnable projection matrix, and b_{proj} is a bias term.
To achieve adaptive attention allocation, a softmax gating mechanism is introduced, described as

$$\alpha_m = \frac{\exp(w_m^T z_m)}{\sum_{m' \in \{t,i,s\}} \exp(w_{m'}^T z_{m'})} \quad (2)$$

where α_m is the gating weight of modality m; $m' \in \{t, i, s\}$ is the modality index, with t denoting text, i denoting image, and s denoting sensor time series; $z_m \in R^{d_h}$ is the encoding vector of modality m; and $w_m \in R^{d_h}$ is a learnable scoring vector for modality m.
The gated fusion representation then follows as

$$h_{\text{gate}} = \sum_{m \in \{t,i,s\}} \alpha_m W_m z_m \quad (3)$$

Where $h_{\text{gate}} \in R^{d_h}$ is the final feature vector after gated weighted fusion, and $w_m \in R^{d_h \times d_h}$ is the modality-specific projection matrix for modality m.

When images are occluded, sensor packets are lost, or text logs are sparse, the gating mechanism automatically reduces the influence of low-quality modalities on the final decision. PEFT achieves parameter-efficient adaptation by freezing the original weights and introducing a low-rank increment shown in Equation 4 to Equation 5, expressed as

$$y = W_0 x + BAx \quad (4)$$

where x is the multimodal fusion input feature formed by concatenating the three modality encoding vectors, y is the true label of the output sample, W_0 is the frozen pretrained weight, A is the rank-reduction projection matrix (input side), and B is the rank-recovery projection matrix (output side).

The KAR module retrieves relevant knowledge from device manuals, historical cases, and safety rules to provide a factual basis for reasoning shown in Equation 6. If the retrieved result set is denoted $\mathcal{D}_{\text{ret}}$, the resulting augmented prompt is

$$\mathcal{P}_{\text{KAR}} = \mathcal{P}_0 \oplus \left[\bigoplus_{d_i \in \mathcal{D}_{\text{ret}}} \text{Format}(d_i) \right] \quad (5)$$

where $\oplus$ denotes context concatenation, $\mathcal{P}_0$ is the base prompt template, $\mathcal{P}_{\text{KAR}}$ is the generated prompt, and $\mathcal{D}_{\text{ret}} \in \{d1, \ldots, dk\}$ is the returned set of top-k relevant documents.

2.2.4. Action execution layer (calibration and Bayesian decision-making)

The final decision follows the Bayesian risk-minimization rule, shown in Equation 6

$$a^*(x) = \underset{a \in \mathcal{A}}{\text{argmin}} [L^T p(x)]_a \quad (6)$$

where $a^*(x)$ is the Bayes-optimal decision output, $p(x)$ is the posterior class probability, L is the cost matrix, and $\mathcal{A}$ is the action space. A mechanism combining a confidence threshold with human review is included to safeguard high-risk scenarios.

2.2.5. Feedback layer (knowledge base and online update)

Execution outcomes and human-review signals are written back to the knowledge base to update the retrieval index and calibration statistics, and, where permitted, to fine-tune the PEFT increments or fusion parameters online, closing the loop for continuous adaptation[21], [22], [23], [24].

This design lets the system balance accuracy, interpretability, and deployment cost at the same time.

2.3. Training Objective and Decision Process

To optimize classification performance, cross-modal consistency, and gating stability, this paper adopts a joint objective function, shown in Equation 7.

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{cls}} + \lambda_1 \mathcal{L}_{\text{align}} + \lambda_2 \mathcal{L}_{\text{gate}} \quad (7)$$

where $\mathcal{L}_{\text{total}}$ is the single scalar optimized by SAGE-IoT; $\mathcal{L}_{\text{cls}}$ is the classification cross-entropy loss; Lalign is a cross-modal alignment loss that constrains the different modalities to remain consistent within a shared semantic space; $\mathcal{L}_{\text{gate}}$ is a gating regularization term that discourages the gating weights from collapsing prematurely onto a single modality; and λ1, λ2 are balancing hyperparameters. The overall decision pipeline can be summarized as "perceptual encoding → gated fusion → knowledge retrieval → semantic reasoning → cost-based decision → feedback update." This preserves the knowledge-organization and reasoning advantages of LLM-style processing while avoiding the resource burden of deploying a full large model on the edge.

3. Experiments and Analysis

The experiments span three representative scenarios: smart-home activity recognition and anomaly detection, industrial rotating-machinery fault warning, and smart-transportation anomaly recognition, corresponding respectively to home security and behavior recognition, equipment health monitoring, and traffic-flow anomaly warning. Each dataset combines text, image, and sensor modalities; results are obtained with 5-fold stratified cross-validation, repeated three times and averaged. To evaluate the proposed framework systematically, six representative baselines spanning three levels, traditional expert systems, classical machine learning, and deep neural networks, were used for comparison:

a. Rule engine: an expert-authored IF–THEN rule set, representing the traditional decision paradigm widely deployed in industrial IoT.
b. Logistic regression: an L2-regularized linear classifier, representing the performance lower bound under linearly separable conditions.
c. Random forest: an ensemble of 50 decision trees (maximum depth 6) combined via bagging.
d. Gradient boosting tree: a boosting ensemble that fits residuals with weak learners over 30 iterations at depth 3, representing the strongest traditional machine-learning baseline.
e. K-nearest neighbors (K = 5): a non-parametric, lazy-learning method that predicts by majority vote among the 5 nearest samples.

f. Multilayer perceptron: a deep neural network with a 128→64 two-hidden-layer structure, ReLU activation, and dropout regularization, used to probe the upper bound of nonlinear mapping for multimodal fusion.

Together, these six baselines span the full spectrum from rule-driven to data-driven methods, giving a comprehensive view of the proposed framework's advantage over different decision-making paradigms. Because the experimental environment does not run a genuine LLM directly, PEFT is approximated with low-rank projection, and KAR is simulated with case retrieval. Beyond accuracy, evaluation also considers average expected cost and the results of component ablation.

2.2. Cross-Scenario Decision Performance Comparison

To systematically assess the effectiveness of the proposed framework on multimodal IoT decision-making tasks, this paper compares it with six baseline methods across three representative scenarios, smart home, industrial rotating machinery, and smart transportation, spanning the full spectrum from rule-driven to data-driven approaches. Table 1 reports the decision accuracy of the proposed method and all baselines across the three scenarios.

Table 1. Results across three scenarios (5-fold cross-validation), unit: %

Method	Smart Home Activity Recognition & Anomaly Detection	Industrial Rotating-Machinery Fault Warning	Smart Transportation Anomaly Recognition
Rule engine	73.5	68.5	70.2
Logistic regression	82	77	79.5
Random forest	85.5	80.5	82.8
Gradient boosting tree	86.8	81.8	84.2
K-nearest neighbors	84	79	81
Multilayer perceptron	88.4	84.2	86.1
Proposed method (SAGE-IoT)	95	88.5	92

As Table 1 shows, the proposed method leads consistently across all three scenarios, with an average accuracy of 91.8%, a gain of 21.1 percentage points over the rule engine, 7.5 points over the gradient boosting tree (the strongest traditional machine-learning baseline), and 5.6 points over the multilayer perceptron (the deep-learning baseline).

More specifically: compared with linear methods such as logistic regression, the proposed method shows a clear advantage on nonlinear multimodal features, confirming that the fusion reasoning mechanism captures cross-modal semantic associations. Compared with ensemble methods such as random forest and gradient boosting trees, its lead is especially pronounced in the industrial scenario, indicating that LLM prior knowledge combined with KAR retrieval augmentation compensates for the limited coverage of rare fault patterns in ensemble learning. Compared with lazy-learning methods such as K-nearest neighbors, end-to-end learning lets the proposed method avoid the distance-metric degradation that occurs in high-dimensional multimodal spaces. Compared with the multilayer perception, its higher average accuracy indicates that PEFT-inspired low-rank projection combined with gated weighted fusion captures deeper cross-modal interactions than nonlinear mapping alone.

Overall, the proposed method outperforms all three baseline categories, rule-driven, traditional machine learning, and deep learning, confirming the effectiveness of jointly designing fusion, retrieval, and lightweight adaptation for multimodal IoT decision-making.

2.3. Modality Ablation and Sample Efficiency

To verify the independent contribution of each component in the proposed framework, two ablation variants were compared; results are shown in Table 2.

Table 2 Ablation results, unit: %

Method	Smart Home Activity Recognition & Anomaly Detection	Industrial Rotating-Machinery Fault Warning	Smart Transportation Anomaly Recognition
Proposed method (full)	95	88.5	92
Proposed method w/o KAR	91.5	84.5	87.8
Proposed method w/o PEFT	93.5	86.8	90.2

Removing the KAR module lowers average accuracy from 91.8% to 87.9%, a drop of 3.9 percentage points that is especially pronounced in the industrial scenario. This indicates that retrieval augmentation contributes substantially to complex anomaly recognition: by introducing the most similar historical cases from the external knowledge base, KAR grounds reasoning in fact and mitigates LLM hallucination on rare fault patterns. Removing the PEFT adapter lowers average accuracy to 90.2%, a drop of 1.6 percentage points, showing that lightweight adaptation, despite introducing only a small number of trainable parameters (0.3% of the total), effectively transfers general language-understanding ability to IoT decision-making tasks.

To examine the independent contribution of each modality, the framework's performance was also tested with a single modality removed at a time, as shown in Table 3.

Table 3 Performance change after removing a single modality, unit: %

Configuration	Accuracy – Home	Accuracy – Industrial	Accuracy – Transportation
Full modalities	86.8	81.8	84.2
Text modality removed	78.2	73.2	75.8
Image modality removed	79.5	74.5	77.2
Sensor modality removed	75.6	69.8	73.8

As Table 3 shows, removing the sensor modality lowers average accuracy by about 11.2 percentage points, removing the text modality lowers it by about 8.6 points, and removing the image modality lowers it by about 7.2 points. All three modalities make an irreplaceable, independent contribution: the sensor modality is most critical for sensing physical anomalies, while the text and image modalities supply semantic context and visual evidence, respectively, forming an effective complement to one another.

To examine the framework's learning efficiency and generalization under scarce labeled data, a sample-ratio comparison was designed for the smart-home scenario. With model structure and hyperparameters fixed, subsets were randomly drawn from the training set at 10%, 20%, 30%, 40%, 60%, 80%, and 100% of its size and trained independently; each subset's resulting model was evaluated for accuracy on the same test set and compared with the rule engine and the gradient boosting tree, as shown in Table 4.

Table 4 Accuracy versus training-data proportion (smart-home scenario), unit: %

Method	10%	20%	30%	40%	60%	80%	100%
Rule engine	68.5	68.5	68.5	68.5	68.5	68.5	68.5
Gradient boosting tree	48.5	55.3	60.1	64.7	73.2	79.5	86.8
Proposed method (SAGE-IoT)	73.5	79.8	83.2	85.2	89.5	93.5	95

4. **Conclusion**

This paper presents a multimodal decision-making framework that brings lightweight LLM adaptation and retrieval-augmented generation into IoT edge decision-making and combines them with the Bayes-optimal decision rule, achieving end-to-end coordination across multimodal perception, knowledge retrieval, lightweight reasoning, and cost-sensitive decision-making. Experiments on three representative scenarios, smart home, industrial rotating machinery, and smart transportation, show that the proposed method reaches an average decision accuracy of 91.8%, improving on the rule engine, gradient boosting tree, and multilayer perceptron by 21.1, 7.5, and 5.6 percentage points, respectively. Ablation experiments confirm the independent contributions of the KAR and PEFT components and show that the sensor, text, and image modalities are strongly complementary; a sample-efficiency analysis shows the framework can approach the full-data performance of traditional methods using only 40% of the training data. More importantly, the cost-sensitive experiments reveal an engineering principle of "calibrate before deciding": feeding uncalibrated probabilities directly into the Bayesian rule instead worsens average expected cost by 594%, so in IoT deployments where false negatives are costly, probability calibration is a necessary precondition for Bayesian decision-making. The framework further achieves parameter compression and compute offloading through low-rank projection and knowledge retrieval, giving it deployment adaptability on resource-constrained edge gateways. Future work will further examine the framework's robustness on real IoT datasets and device pipelines, and systematically compare how different retrieval granularities, adaptation ranks, and calibration strategies affect decision quality.

References

[1] A. Giuliano, A. McCafferty-Leroux, J. Yawney, and S. Andrew Gadsden, "Cognitive Internet of Things: A Review of Theory, Applications, and Recent Advances," *IEEE Commun. Surv. Tutor.*, vol. 28, pp. 446–484, 2026, doi: 10.1109/COMST.2025.3615461.

[2] T. Sadiq and C. W. Omlin, "Sensing in Smart Cities: A Multimodal Machine Learning Perspective," *Smart Cities*, vol. 9, no. 1, p. 3, Jan. 2026, doi: 10.3390/smartcities9010003.

[3] K. Taha, "Big Data Analytics in IoT, social media, NLP, and information security: trends, challenges, and applications," *J. Big Data*, vol. 12, no. 1, p. 150, Jun. 2025, doi: 10.1186/s40537-025-01192-9.

[4] O. B. Sezer, E. Dogdu, and A. M. Ozbayoglu, "Context-Aware Computing, Learning, and Big Data in Internet of Things: A Survey," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 1–27, Feb. 2018, doi: 10.1109/JIOT.2017.2773600.

[5] G. Qu, Q. Chen, W. Wei, Z. Lin, X. Chen, and K. Huang, "Mobile Edge Intelligence for Large Language Models: A Contemporary Survey," *IEEE Commun. Surv. Tutor.*, vol. 27, no. 6, pp. 3820–3860, Dec. 2025, doi: 10.1109/COMST.2025.3527641.

[6] A. Sharshar, L. U. Khan, W. Ullah, and M. Guizani, "Vision-Language Models for Edge Networks: A Comprehensive Survey," *IEEE Internet Things J.*, vol. 12, no. 16, pp. 32701–32724, Aug. 2025, doi: 10.1109/JIOT.2025.3579032.

[7] J.-S. Yang, Z. Shen, Z. Zeng, and Z. Chen, "Domain-Adapted Large Language Models for Industrial Applications: From Fine-Tuning to Real-Time Deployment," *Comput. Sci. Bull.*, vol. 8, no. 01, pp. 272–289, Dec. 2025, doi: 10.71465/csb162.

[8] F. Sarhaddi *et al.*, "LLMs and IoT: A Comprehensive Survey on Large Language Models and the Internet of Things," *IEEE Open J. Commun. Soc.*, vol. 7, pp. 3585–3616, 2026, doi: 10.1109/OJCOMS.2026.3680064.

[9] H. Xu, L. Han, Q. Yang, and M. Li, "Penetrative AI: Making LLMs Comprehend the Physical World | Proceedings of the 25th International Workshop on Mobile Computing Systems and Applications," ACM Conferences. Accessed: Sep. 10, 2026. [Online]. Available: https://dl.acm.org/doi/10.1145/3638550.3641130

[10] A. Liu, W. Jiang, and Z. Feng"Multi-Modal Integrated Sensing and Communication in Internet of Things With Large Language Models." Accessed: Sep. 10, 2026. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11026879

[11] T. An, Y. Zhou, H. Zou, and J. Yang, "IoT-LLM: A framework for enhancing large language model reasoning from real-world sensor data," *Patterns*, vol. 7, no. 1, Jan. 2026, doi: 10.1016/j.patter.2025.101429.

[12] D. Jiang, Z. Shen, Q. Zheng and J. Jin"Farm-LightSeek: An Edge-Centric Multimodal Agricultural IoT Data Analytics Framework With Lightweight LLMs." Accessed: Sep. 10, 2026. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11027609

[13] M. K, R. G, A. Mondal, S. Singh, S. Tiwari, and J. Sharma, "HARMONY: A Framework for Multimodal LLM-Powered AI Agents in Smart Homes via the Model Context Protocol," *IEEE Access*, vol. 14, pp. 13669–13687, 2026, doi: 10.1109/ACCESS.2026.3653992.

[14] H. Zhang, M. Farzanullah, M. Ghassemi, A. Bin Sediq, A. Afana, and M. Erol-Kantarci, "Multi-Modal Data-Enhanced Foundation Models for Prediction and Control in Wireless Networks: A Survey," *IEEE Commun. Surv. Tutor.*, vol. 28, pp. 4359–4393, 2026, doi: 10.1109/COMST.2025.3648785.

[15] J. Li, J. Li, G.. Yang, H. Chi and L. Yang“Applications of Large Language Models and Multimodal Large Models in Autonomous Driving: A Comprehensive Review.” Accessed: Sep. 10, 2026. [Online]. Available: https://www.mdpi.com/2504-446X/9/4/238

[16] C. Zhang, Z. Yang, X. He, and L. Deng, “Multimodal Intelligence: Representation Learning, Information Fusion, and Applications,” *IEEE J. Sel. Top. Signal Process.*, vol. 14, no. 3, pp. 478–493, Mar. 2020, doi: 10.1109/JSTSP.2020.2987728.

[17] W. Shao, D. Fan, C. Cui, Y. Xu, and X. Lyu“Deep Multimodal Data Fusion,” *ACM Comput. Surv.*, Accessed: Sep. 10, 2026. [Online]. Available: https://dl.acm.org/doi/10.1145/3649447

[18] P. Wang, S. Song, H. Ji, S. Cao, and J. Hu“From Models to Systems: A Comprehensive Survey of Efficient Multimodal Learning | OpenReview.” Accessed: Sep. 10, 2026. [Online]. Available: https://openreview.net/forum?id=yfTU8FTS2Z

[19] S. Luo *et al.*, “Delving Into Multi-Modal Multi-Task Foundation Models for Road Scene Understanding: From Learning Paradigm Perspectives,” *IEEE Trans. Intell. Veh.*, vol. 9, no. 12, pp. 8040–8063, Dec. 2024, doi: 10.1109/TIV.2024.3406372.

[20] T. Jiao, C. Guo, X. Feng, Y. Chen, and J. Song, “A Comprehensive Survey on Deep Learning Multi-Modal Fusion: Methods, Technologies and Applications,” *Comput. Mater. Contin.*, vol. 80, no. 1, pp. 1–35, Jul. 2024, doi: 10.32604/cmc.2024.053204.

[21] S. Lim, H.-C. Kwon, and M. Kim, “When to Invoke an LLM in Industrial Natural-Language-to-DSL Conversion: A Taxonomy-Driven, Confidence-Gated Selective Hybrid for Safety-Critical Avionics Test-Script Authoring,” *IEEE Access*, pp. 1–1, 2026, doi: 10.1109/ACCESS.2026.3730474.

[22] C. Xu *et al.*, “LiveFact: A Dynamic, Time-Aware Benchmark for LLM-Driven Fake News Detection,” in *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, M. Liakata, V. P. Moreira, J. Zhang, and D. Jurgens, Eds., San Diego, California, United States: Association for Computational Linguistics, Jul. 2026, pp. 11881–11910. doi: 10.18653/v1/2026.acl-long.546.

[23] C. Xu *et al.*, “LiveFact: A Dynamic, Time-Aware Benchmark for LLM-Driven Fake News Detection,” in *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, M. Liakata, V. P. Moreira, J. Zhang, and D. Jurgens, Eds., San Diego, California, United States: Association for Computational Linguistics, Jul. 2026, pp. 11881–11910. doi: 10.18653/v1/2026.acl-long.546.

[24] F. Yang, X. Zhang and Z. Yu“Human–Machine Collaborative Learning for Streaming Data-Driven Scenarios.” Accessed: Sep. 10, 2026. [Online]. Available: https://www.mdpi.com/1424-8220/25/21/6505