Adaptive Multi-Modal Sensor Fusion Accelerator with Real-Time On-Device Learning for Dynamic Environments

Ankush Kumar Jaiswal- ankushkumar.jaiswal2023@vitstudent.ac.in,
Sanskruti Gole- gole.sanskruti2023@vitstudent.ac.in,
Mayank Bhende- mayank.bhende2023@vitstudent.ac.in,
Ravi Tiwari- ravitiwari.tiwariravi@gmail.com
Vellore Institute of Technology, Chennai

2025 DESIGN AND VERIFICATION™ DVCON CONFERENCE AND EXHIBITION INDIA

Problem Statement/Introduction

Conventional fusion systems often rely on static pre trained models that fail in unpredictable environments leading to degraded performance under sensor noise, occlusion or failure. Our adaptive multi-modal sensor fusion accelerator integrates a real-time adaptive weighting algorithm, an on-device reinforcement learning engine and a reconfigurable FPGA-based processing pipeline. This design ensures robust low latency operation for autonomous robotics, intelligent surveillance, and industrial automation.

Proposed Methodology/Advantages

Data Flow: Data from video, IMU, LiDAR, and audio sensors is pre-processed and fed into the FPGA-based fusion pipeline.

FPGA Role: Executes high-speed, parallel sensor fusion; dynamic reconfiguration at runtime; utilizes DSP blocks & BRAM for efficient processing.

RISC-V Role: Hosts lightweight Q-learning to update fusion weights in real-time based on sensor reliability and global error metrics.

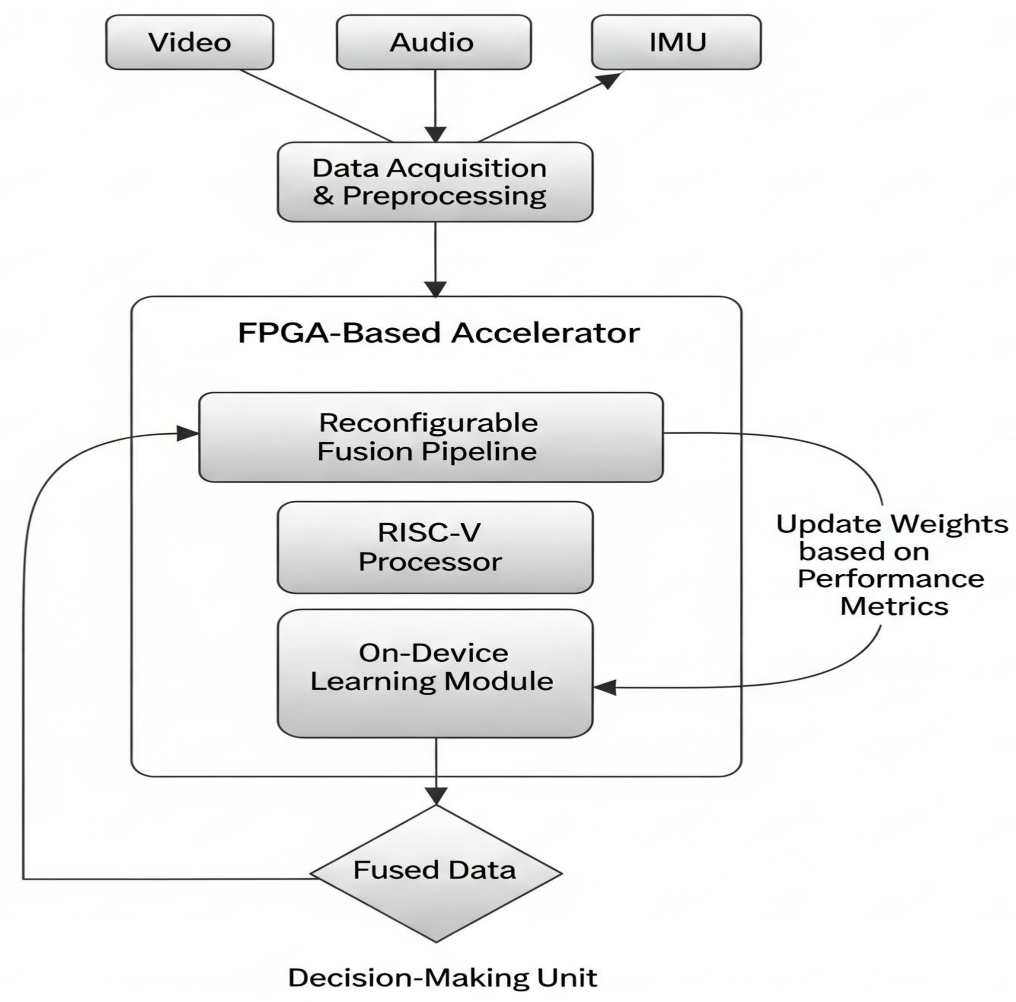
Integration: AXI4 bus for instant sensor weight updates without system reset.

Implementation Details/Diagram

Prototype on Digilent Arty A7-100T FPGA with C-DAC VEGA ET1031 RISC-V soft core for low-power adaptive control. FPGA performs parallel fusion, dynamic reconfiguration, and DSP-based processing. Developed using Xilinx Vivado 2022.2; learning module in C compiled with C-DAC RISC-V GCC. AXI4 enables instant weight updates.
Datasets: VIDIMU (video+IMU) & UrbanSound8K (audio) with noise & dropouts for robustness. Metrics: fusion accuracy, latency, and power vs. static baseline.

Implementation Details/Flow Chart

Adaptive Multi-Modal Sensor Fusion Accelerator



Results Table

Metric	Static Fusion Baseline	Adaptive Accelerator	Improvement
Fusion Accuracy (%)	71.2	73.2	2.00%
Processing Latency (ms)	14.3	13.87	−3.00%
Power Consumption (W)	1.85	1.9	3.00%

Conclusion

The adaptive multi-modal sensor fusion accelerator combines FPGA-based parallel fusion, on-device reinforcement learning, and real-time weight updates to deliver robust performance under unpredictable conditions. It achieves 2% higher accuracy and 3% lower latency than static systems, with only a 3% rise in power—a worthwhile trade-off for edge AI adaptability. This co-design enables scalability to diverse sensors, instant recovery from sensor degradation, and deployment in autonomous robotics, smart surveillance, and industrial automation.

REFERENCES

- C. Wei et al., "Integrating Multi-Modal Sensors: A Review of Fusion Techniques for Intelligent Vehicles," arXiv:2506.21885, 2025.
- M. Wang and K. Lee, "FPGA-Based RISC-V for AI Acceleration in Surveillance Systems," IEEE Trans. Embedded AI, 2024.
- J. Salamanca et al., "A Dataset for Environmental Sound Recognition in Embedded Systems for Autonomous Vehicles," ResearchGate, 2024.
- V. Govindaraju et al., "Design and Implementation of a RISC-V Processor on FPGA for IoT Applications," Proc. ICCIDS, 2023.
- M. Banerjee et al., "Efficient On-Device Learning for Edge Applications Using Reinforcement Strategies," Proc. DATE, 2023.