

Evaluating Multimodal Fusion Strategies for Resilient Agricultural Sensing Systems

Guide Name	Panel Head
Dr. Sornalakshmi K	Dr. Sornalakshmi K

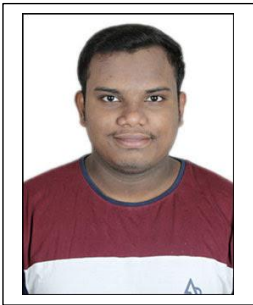
Faculty Advisor	Project Domain
Dr. K ShanthaKumari	

Student(s) Details: Name

1. Abhay Shaji Valiyaparambil

2. Ponnuri Aniruddha

Passport size photo(s)



Registration Number(s)

1. RA2112704010006

2. RA2112704010015

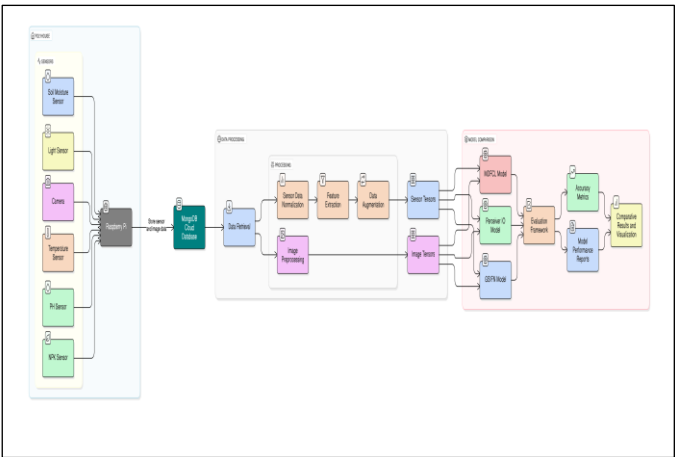
Email ID(s)&Mobile Number(s)

1: as3735@srmist.edu.in , 9104909364	2: pp0783@srmist.edu.in , 9205481551
--	--

Abstract

Precision agriculture increasingly relies on integrating heterogeneous sensor streams—such as time series data from soil sensors (measuring moisture, pH, NPK) and high-resolution crop images—to enhance predictive analytics and crop management. However, fusing these diverse data types presents significant challenges due to differences in temporal resolution, modality, and data quality. In this study, we systematically compare three state-of-the-art multimodal fusion networks—MDFCL, GSIFN, and Perceiver IO—on a unique dataset of synchronized soil sensor readings and field images collected under varying conditions. Each fusion approach encapsulates a distinct strategy: MDFCL constructs modality-specific graphs and aligns their representations through contrastive learning, promoting robustness to sensor corruption. GSIFN interlaces modality-specific masks within a unified Transformer, facilitating strong cross-modal interactions and resilience to partial image loss. Perceiver IO employs an asymmetric attention bottleneck to efficiently compress heterogeneous inputs, enabling scalable, real-time inference with quasi-linear complexity. Empirical results show that all three models outperform unimodal baselines in predicting agronomic traits, but differ in robustness and computational efficiency. MDFCL is most stable under sensor dropout, GSIFN excels with incomplete imagery, and Perceiver IO balances accuracy with scalability. These findings provide practical guidance for deploying robust sensor and image fusion solutions in precision agriculture, and offer a template for multimodal data integration in other domains

Architecture Diagram



Significance of the Project

This project is significant because it leverages IoT-enabled multimodal data fusion to transform precision agriculture by integrating high-frequency sensor data with crop imagery. This integration enables more accurate, real-time monitoring of crop health and nutrient levels, improving decision-making and resource management in controlled environments. By systematically comparing advanced fusion models, the project identifies robust, scalable solutions suitable for resource-constrained settings, enhancing both productivity and sustainability. Ultimately, it supports optimized input use, reduces environmental impact, and empowers farmers with data-driven insights, contributing to increased yields, cost savings, and sustainable agricultural practices.

Conclusion

This study provides a rigorous comparison of three leading multimodal fusion architectures—GSIFN, MDFCL, and Perceiver IO—for soil potassium classification using synchronized sensor and image data. While all models achieved near-perfect accuracy under clean conditions, each exhibited unique strengths: GSIFN excelled in rapid convergence and stability, MDFCL demonstrated superior robustness to sensor corruption, and Perceiver IO offered a strong balance between accuracy and computational efficiency. These findings highlight that model selection in precision agriculture should be guided not just by peak accuracy, but by practical deployment factors such as sensor reliability, computational resources, and real-time adaptation needs.

Conference/Journal Publication Details (Mandatory)

6th International Conference on Data Science and Applications.

Track Name: ICDSA2025

Paper ID: 1166

Paper Title: Evaluating Multimodal Fusion Strategies for Resilient Agricultural Sensing Systems