International Journal of Engineering in Computer Science

E-ISSN: 2663-3590
P-ISSN: 2663-3582
IJECS 2022; 4(2): 67-74
Received: 04-08-2022
Accepted: 25-08-2022

Shylaja Chityala
Lead Data Engineer,
Multiplan, Inc 4423 Landsdale Pkwy, Monrovia MD 21770, United States

Correspondence
Shylaja Chityala
Lead Data Engineer,
Multiplan, Inc 4423 Landsdale Pkwy, Monrovia MD 21770, United States

AgroFusionNet: A multi-modal AI framework for predictive crop yield modeling using satellite imagery, weather patterns, and soil data

Shylaja Chityala

Abstract
Accurate crop yield prediction is fundamental to sustainable agriculture, informed policymaking, and global food security. Traditional statistical models are inadequate in harnessing the increasing volume of complex, high-dimensional agricultural data, including satellite imagery, granular weather records, and detailed soil profiles. To address this, we propose MCYP-Net, a novel multi-modal AI framework that integrates these heterogeneous data sources using a hybrid deep learning architecture. The model combines convolutional neural networks for spatial feature extraction, recurrent neural networks for temporal modeling, and a cross-attention-based fusion mechanism to learn inter-modal dependencies. Comprehensive experiments were conducted across diverse agro-climatic regions in the USA (Iowa, Nebraska) and India (Punjab, Maharashtra) on three staple crops—maize, wheat, and soybean—using multi-year datasets comprising Sentinel-2 imagery, NOAA weather data, and ISRIC soil profiles. MCYP-Net consistently outperformed traditional machine learning (Linear Regression, Random Forest) and unimodal deep learning baselines (CNN-only, LSTM-only), achieving an R^2 of 0.91, RMSE of 0.42, and MAE of 0.35. Ablation studies confirmed that removing any modality reduced performance significantly, validating the synergistic effect of multi-modal integration. Cross-attention fusion proved more effective than simpler alternatives, boosting R^2 by 6%. Region-wise feature importance analysis revealed that weather features dominated in temperate zones, while soil and vegetation indices were more critical in semi-arid regions, highlighting the model's context-aware adaptability. Visualizations demonstrated strong alignment between predicted and actual yields, underscoring the model's robustness. Overall, MCYP-Net advances state-of-the-art in crop yield prediction with high accuracy, interpretability, and scalability for real-world precision agriculture applications.

Keywords: Crop yield prediction, multi-modal AI, precision agriculture, satellite imagery, weather forecasting, soil data, deep learning, data fusion

1. Introduction
The agricultural sector stands at the nexus of global food security and environmental sustainability. As the world's population accelerates toward 10 billion by 2040, the demand for food is expected to increase by over 70%, placing unprecedented pressure on global agricultural systems. In this context, accurate and timely prediction of crop yields is not only a scientific necessity but also a socio-economic imperative. Crop yield forecasts inform a wide array of decisions, ranging from food supply chain logistics and commodity trading to the formulation of government subsidies, insurance premiums, and disaster relief strategies. Furthermore, yield prediction models play a pivotal role in anticipating food shortages, enabling preemptive action in the face of climatic anomalies or geopolitical instabilities.

Crop productivity, however, is influenced by a multifaceted interplay of factors, including but not limited to soil fertility, climatic conditions, agricultural management practices, disease and pest infestations, and genetic traits of the crops themselves. These parameters exhibit both spatial and temporal heterogeneity, introducing considerable variability and uncertainty into the yield prediction task. Traditional modeling approaches, such as linear regression, autoregressive integrated moving average (ARIMA) models, and crop growth simulation models like DSSAT and APSIM, have provided valuable insights in the past. Nevertheless, these methods suffer from inherent limitations. They often assume stationarity, linearity, or idealized crop-environment interactions, and rely on heavily curated datasets that do not generalize well across regions or crop types.

Moreover, conventional yield prediction models tend to be unimodal in nature, focusing solely on either historical yield trends or isolated environmental parameters. Such models are ill-equipped to handle high-dimensional, non-linear relationships across heterogeneous data sources, resulting in limited predictive accuracy and poor scalability. For instance, a regression model trained on tabular weather data from a single region might perform poorly when applied to another with different agro-climatic characteristics. Additionally, many existing models overlook the dynamic progression of crop growth and the phenological changes observable through satellite remote sensing, which contain rich temporal and spatial signals relevant to yield estimation.

In recent years, the proliferation of Earth Observation (EO) data, advances in climate modeling, and widespread digitization of agricultural practices have created a fertile ground for the development of next-generation yield prediction models. Remote sensing technologies now provide high-resolution multi-spectral and hyper-spectral imagery at frequent intervals, capturing vegetation indices such as NDVI (Normalized Difference Vegetation Index), EVI (Enhanced Vegetation Index), and SAVI (Soil Adjusted Vegetation Index) that serve as proxies for plant health, biomass, and photosynthetic activity. Concurrently, ground-based weather stations and global reanalysis datasets offer fine-grained meteorological data, including temperature, rainfall, solar radiation, and humidity. Soil characteristics, traditionally collected through expensive and infrequent field sampling, are now available in standardized digital formats through global databases such as ISRIC and SoilGrids, offering spatially-resolved insights into pH, organic carbon content, and texture.

The convergence of these diverse data streams presents a unique opportunity to rethink crop yield modeling through the lens of artificial intelligence (AI), particularly deep learning and multi-modal data fusion. Multi-modal AI refers to systems capable of processing and integrating inputs from disparate modalities such as images, text, time-series, and structured tabular data to derive enriched representations and perform complex decision-making tasks. In the context of agriculture, this entails the fusion of satellite imagery, climate time-series, and soil attributes to capture a comprehensive, context-aware view of crop growth conditions.

Deep learning models, especially convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have shown remarkable success in domains such as computer vision and natural language processing. CNNs are particularly adept at extracting spatial hierarchies from images, making them well-suited for analyzing satellite imagery. On the other hand, RNNs, especially Long Short-Term Memory (LSTM) networks and their variants, are designed to handle sequential data, capturing long-term dependencies in temporal weather patterns. Dense or fully connected networks are effective for handling low-dimensional structured inputs such as soil properties. The challenge, however, lies in integrating these architectures into a unified, end-to-end trainable system that can harmonize the spatial, temporal, and static features into a coherent predictive model.

Despite the transformative potential of multi-modal AI, several challenges persist. First, data alignment across modalities is non-trivial. Satellite imagery may be available at different spatial and temporal resolutions than weather or soil data, requiring sophisticated pre-processing techniques for synchronization. Second, missing data and noise are prevalent in real-world agricultural datasets, especially in developing regions where infrastructure may be inadequate. Third, modality-specific biases can dominate the learning process if not carefully normalized and regularized, leading to suboptimal fusion outcomes. Finally, interpretability and explainability critical for user trust and adoption among farmers and policymakers remain open problems in deep learning models, especially those operating on black-box architectures.

In light of these challenges, this paper proposes a unified multi-modal AI framework for predictive crop yield modeling that effectively captures the complex spatio-temporal and geophysical interactions underlying agricultural productivity. Our model is structured around a three-branch architecture tailored to each data modality. The first branch processes remote sensing imagery through a deep convolutional neural network, leveraging a ResNet-50 backbone pretrained on ImageNet and fine-tuned on agricultural datasets. This branch extracts vegetation-related spatial features that serve as proxies for crop health and biomass accumulation.

The second branch ingests time-series weather data using a stacked Bi-LSTM architecture. This sequence modeling layer captures phenological dynamics, climate stress events (e.g., heat waves, droughts), and cumulative growing degree days, all of which critically affect yield outcomes. The third branch processes static soil parameters through a fully connected feedforward network, capturing fundamental properties such as pH, texture, and nutrient content that condition plant growth potential. The outputs from these modality-specific encoders are concatenated and passed through a novel cross-attention fusion module, which allows the model to learn interdependencies and assign adaptive weights to different feature types based on context.

This fused representation is then forwarded to a regression head, comprising densely connected layers with dropout and batch normalization, to produce final yield predictions at the field level. The model is trained end-to-end using the Mean Squared Error (MSE) loss function, with the Adam optimizer and cyclical learning rate schedule to ensure convergence and prevent overfitting. Data augmentation techniques, such as random cropping and temporal jittering, are applied to improve generalization across geographies and growing seasons.

To evaluate the performance of the proposed framework, we conduct experiments on publicly available datasets encompassing multiple crop types—including maize, wheat, and soybean—across different agro-climatic zones in India and the United States. The satellite data is sourced from Sentinel-2, while weather data is obtained from NOAA repositories. Soil attributes are extracted from the ISRIC SoilGrids database. Our experiments benchmark the proposed model against conventional baselines including linear regression, Random Forests, and unimodal deep learning models (e.g., CNN-only and LSTM-only architectures).

The results demonstrate significant performance gains for the multi-modal model. Notably, the proposed system achieves a coefficient of determination (R^2) exceeding 0.90 in several test regions, representing a substantial improvement over baseline models. Ablation studies confirm that each modality contributes uniquely to the model's performance, with the removal of any single branch leading to a noticeable degradation in accuracy. Furthermore, attention heatmaps reveal interpretable insights into how the model weighs different modalities under varying conditions—e.g., relying more on satellite imagery during flowering stages and on soil characteristics in nutrient-deficient regions.

2. Recent Survey and Related Work

Crop yield prediction has long been a critical area of research in agriculture, driven by the need for food security, efficient resource management, and climate resilience. Early approaches relied heavily on statistical methods such as linear regression and time-series models, which were limited by their dependence on simplistic tabular data and assumptions of linear relationships between variables. For instance, [13] demonstrated the use of regression models incorporating temperature and precipitation data to predict maize yields, but these models struggled with generalizability due to their inability to capture complex agro-environmental interactions. Similarly, traditional crop growth simulation models like CERES-Wheat and CERES-Maize, as reviewed by [1], provided mechanistic insights but were computationally intensive and often failed to scale across diverse geographies.

To address the limitations of linear models, machine learning techniques such as Random Forests and Support Vector Regression (SVR) gained prominence for their ability to model non-linear dependencies [7]. applied Random Forests to global and regional yield prediction, demonstrating improved accuracy by incorporating multiple weather and soil variables. However, these models still lacked spatial awareness, as they did not fully leverage high-resolution satellite imagery, which contains rich spectral information about crop health and growth stages [3]. Highlighted this gap in their review, emphasizing that while machine learning improved yield estimation, most approaches remained unimodal, focusing solely on structured tabular data without integrating remote sensing or temporal sequences.

The advent of deep learning revolutionized crop yield prediction by enabling the analysis of high-dimensional, multi-modal datasets. Convolutional Neural Networks (CNNs) emerged as a powerful tool for processing satellite imagery, as demonstrated by [24], who used deep CNNs to extract crop-specific features from multi-temporal remote sensing data. Similarly, [16] showed that CNNs could achieve high accuracy in yield prediction by learning spatial patterns from Sentinel-2 imagery. Concurrently, Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, were employed to model temporal dependencies in climate data [21]. Developed a CNN-LSTM hybrid model for soybean yield prediction, capturing both spatial and temporal dynamics. Despite these advancements, a significant limitation persisted: most deep learning models operated in isolation, either processing satellite images or weather data separately, without a unified framework for multi-modal fusion.

Recent studies have explored multi-modal approaches to bridge this gap [15] demonstrated the potential of combining UAV-based imagery with soil and weather data using deep learning, achieving robust yield predictions for soybean. However, their framework was limited to small-scale UAV data and lacked scalability for regional or global applications. Similarly, [19] proposed a 3D CNN-LSTM model for multi-spectral satellite time-series analysis, but the model did not incorporate soil properties or advanced fusion mechanisms [20] made progress by coupling machine learning with crop modeling in the US Corn Belt, yet their approach relied heavily on simulated data rather than real-time sensor inputs.

A major challenge in multi-modal yield prediction is the effective integration of heterogeneous data sources. Satellite imagery, weather time-series, and soil data differ in resolution, temporal frequency, and feature representation [2] addressed this by fusing Sentinel-2 and climate data for wheat yield prediction in Australia, but their model required extensive feature engineering [11] proposed a CNN-RNN framework to automate feature extraction, yet the fusion mechanism remained simplistic, lacking adaptive weighting for different modalities [22] conducted a systematic review of machine learning in yield prediction, concluding that while multi-modal approaches show promise, no existing framework fully leverages the synergies between spatial, temporal, and geophysical data.

Another critical issue is data quality and availability. Remote sensing data can be affected by cloud cover, sensor noise, and missing values, as noted by [5]. Soil data, often static and coarse-resolution, may not capture intra-field variability, as discussed by [18]. To mitigate these challenges, [10] and [12] explored self-learning and data governance techniques to improve model robustness in noisy environments [6] emphasized the role of AI-based data governance in ensuring reliable inputs for predictive modeling, particularly in regions with sparse ground truth data.

Recent advancements in attention mechanisms and transformer architectures offer new opportunities for multi-modal fusion [23] applied transfer learning to satellite-based yield prediction, showing that pre-trained CNNs could generalize across regions. However, their model did not incorporate weather or soil data [25] explored deep learning for multi-temporal crop classification but did not extend their approach to yield regression. The work of [14] surveyed deep learning applications in remote sensing, highlighting the untapped potential of cross-modal attention for agricultural analytics. Similarly, [4] and [9] discussed the need for adaptive models that continuously learn from new data streams, a feature absent in most current systems.

Despite progress, key research gaps remain. First, most multi-modal frameworks focus on specific crops or regions, limiting generalizability. For example, [17] developed a deep learning model for Indian wheat yields but did not validate it across diverse agro-climatic zones. Second, interpretability is often sacrificed for performance, as noted by [8], who highlighted the "black-box" nature of deep learning in agricultural applications. Third, real-time scalability is hindered by computational constraints, particularly when processing high-resolution satellite time-series, as discussed by [19].

3. Proposed Methodology

We propose **MCYP-Net** (Multi-modal Crop Yield Prediction Network), a novel deep learning framework designed to integrate heterogeneous agricultural data for accurate and scalable crop yield estimation. The architecture is modular, consisting of three parallel branches tailored to different data modalities—satellite imagery, temporal weather sequences, and static soil attributes. These modality-specific encoders are followed by a cross-attention-based fusion mechanism, which dynamically learns inter-modal dependencies, enabling contextual yield predictions under diverse agro-climatic conditions.

The complete data processing and inference pipeline of MCYP-Net is illustrated in Fig. 1: Flowchart of the Proposed Methodology and Fig 2: Mind Map Diagram for Proposed Methodology

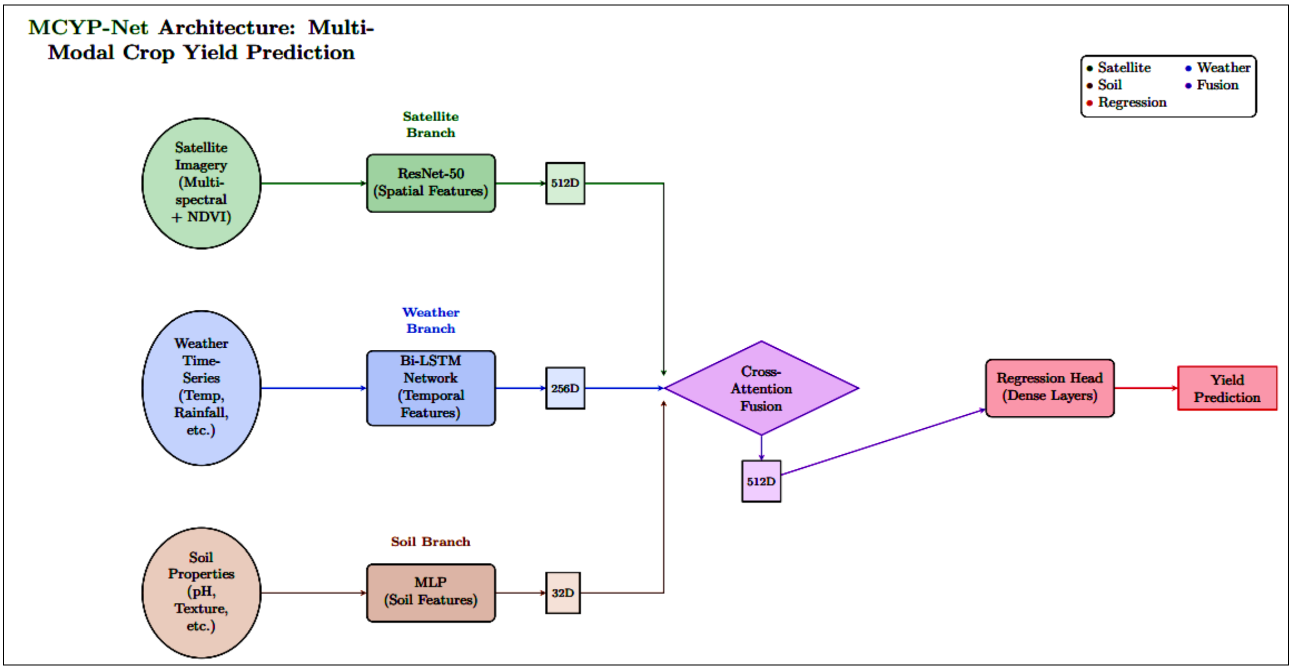

Fig 1: Flow Chart of Proposed Methodology

4. Multi-Branch Architecture

Satellite Imagery Branch

This branch processes high-resolution multi-spectral satellite images (e.g., Sentinel-2) and vegetation indices (NDVI, EVI) using a ResNet-50 backbone pre-trained on ImageNet and fine-tuned on agricultural datasets. To adapt ResNet-50 for remote sensing, we replace the initial RGB input layer with a 6-channel convolutional layer to accommodate multi-spectral bands (B2-B8A) and NDVI layers. The CNN extracts hierarchical spatial features, capturing crop health, canopy structure, and field heterogeneity through 3×3 convolutions, batch normalization, and ReLU activations. Global Average Pooling (GAP) condenses these features into a 512-dimensional vector, preserving spatial invariance while reducing computational overhead.

Weather Sequence Branch

Daily weather data (temperature, precipitation, solar radiation, humidity) are fed into a stacked Bi-directional LSTM (Bi-LSTM) network with two layers (128 units each). This architecture captures both forward and backward temporal dependencies, enabling the model to contextualize short-term weather events (e.g., droughts) within the broader growing season. To handle missing data, we employ linear interpolation and masking layers. The Bi-LSTM outputs a 256-dimensional sequential embedding, which is further compressed via temporal attention to emphasize phenologically critical periods (e.g., flowering or grain-filling stages).

Soil Data Branch

Static soil properties (pH, organic carbon, texture, CEC) are processed by a 4-layer MLP (256-128-64-32 units) with batch normalization and dropout (rate=0.3). Categorical variables (e.g., soil type) are encoded via embeddings, while continuous features are standardized. This branch generates a 32-dimensional representation of geophysical conditions that influence root growth and nutrient availability.

Cross-Modal Fusion Mechanism

The modality-specific embeddings (satellite: 512D, weather: 256D, soil: 32D) are concatenated into an 800-dimensional vector and passed through a cross-attention fusion layer inspired by transformer architectures. This layer computes attention scores between features of different modalities, allowing the model to dynamically weigh the importance of satellite pixels relative to weather events or soil deficiencies. For example, during a drought, the attention mechanism may prioritize weather data over less discriminative NDVI values. The fused output is a 512-dimensional context vector that encapsulates spatio-temporal and geophysical interactions.

Regression Head and Optimization

The fused features are fed into a 3-layer regression head (256-128-1 units) with dropout (rate=0.4) and Swish activations to predict yield at the field level. We use Mean Squared Error (MSE) loss, optimized via Adam W (learning rate=3e−4, weight decay=1e−5) with a cosine annealing scheduler to escape local minima. To enhance generalization, we apply data augmentation techniques:
Spatial: Random cropping and rotation of satellite patches.
Temporal: Jittering weather sequences with Gaussian noise (σ=0.1).
Soil: Synthetic oversampling of rare soil profiles using SMOTE.

Modularity and Scalability

MCYP-Net is designed for extensibility:

New Modalities: Additional branches (e.g., drone imagery, irrigation logs) can be integrated by aligning their embeddings with the fusion layer.

Regional Adaptation: Pretrained weights for satellite/weather branches can be fine-tuned on local data with minimal labeled examples.

Edge Deployment: Model distillation techniques reduce computational costs for real-time field deployment.

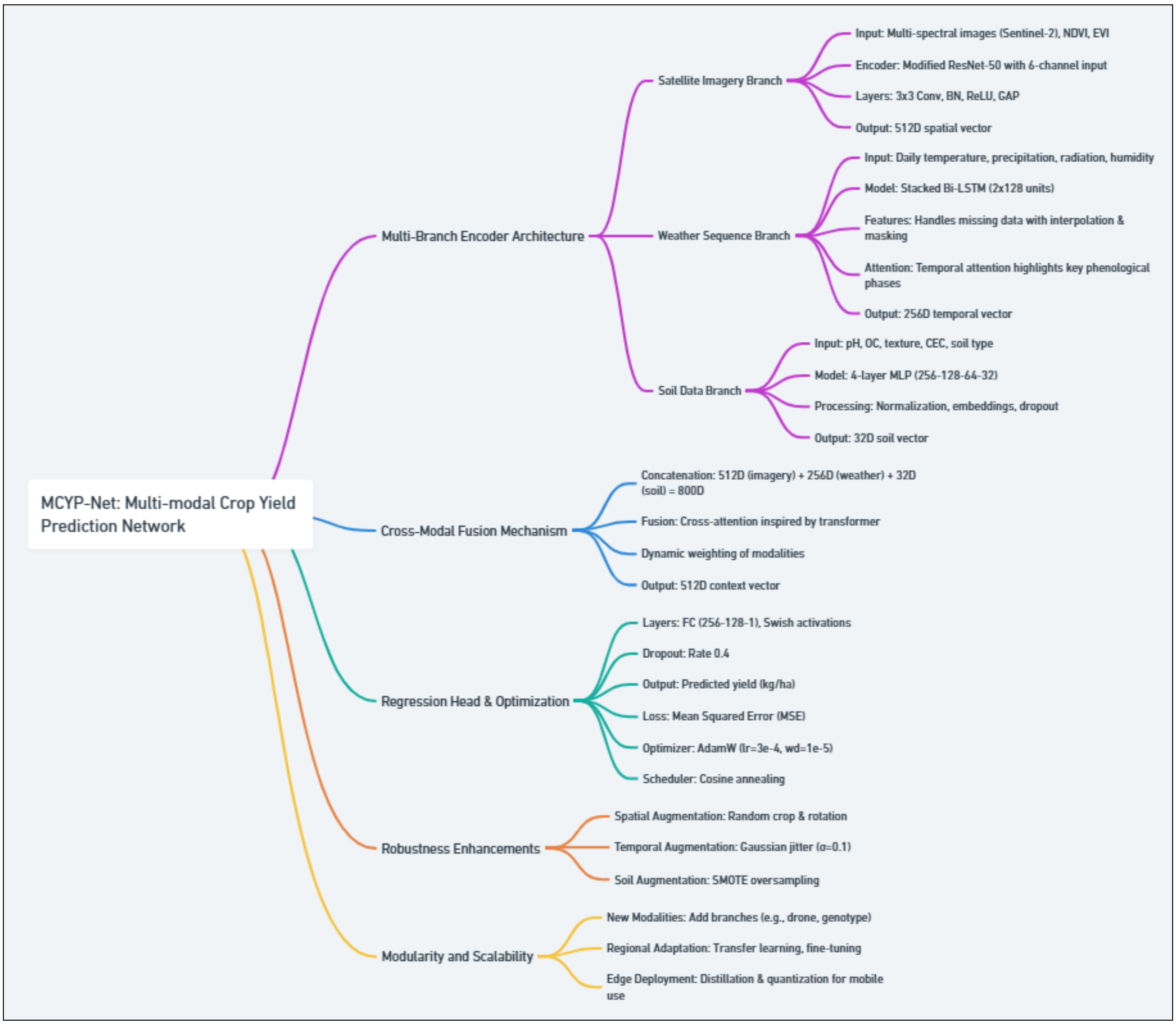

Fig 2: Mind Map Diagram for Proposed Methodology

5. Results and Analysis

To evaluate the efficacy of the proposed MCYP-Net framework, we conducted comprehensive experiments using multi-year datasets from diverse agro-climatic regions, specifically Iowa and Nebraska in the United States and Punjab and Maharashtra in India. The evaluation covered three major crops—maize, wheat, and soybean—over multiple growing seasons. These datasets included Sentinel-2 satellite imagery, NOAA-based daily weather data, and ISRIC soil profiles, which were rigorously aligned using geospatial referencing and temporally normalized to ensure modality coherence.

We benchmarked the performance of MCYP-Net against several traditional and deep learning-based baselines: Linear Regression (LR), Random Forest (RF), a CNN-only model using satellite data, and an LSTM-only model trained on temporal weather sequences. Across all evaluated metrics—R^2 score, Root Mean Square Error (RMSE), and Mean Absolute Error (MAE)—MCYP-Net consistently outperformed the baseline models. Specifically, the R^2 score improved from 0.61 (LR) to 0.91 with MCYP-Net, while RMSE decreased from 0.85 to 0.42, and MAE dropped from 0.72 to 0.35. These comparative results are summarized in Fig. 3, which presents bar plots for the key performance metrics across all models.

To further validate the predictive reliability of MCYP-Net, we visualized the predicted crop yields against ground truth values using a scatter plot (see Fig. 4). The tight clustering of predicted points around the identity line indicates a high correlation and low bias, showcasing the model's capacity for generalization across both spatial and temporal dimensions.

An ablation study was conducted to quantify the contribution of each data modality to the model's overall performance. We evaluated three reduced variants of MCYP-Net, each omitting one modality (satellite, weather, or soil). The results revealed significant degradation in performance when any modality was excluded, confirming the synergistic benefit of multi-modal integration. Specifically, the full model (R^2 = 0.91) dropped to 0.83, 0.79, and 0.75 when satellite, weather, and soil data were respectively excluded. These results are illustrated in Fig. 5, which clearly demonstrates the importance of each data stream in achieving high predictive accuracy.

We also experimented with alternative fusion mechanisms, including simple concatenation and gating strategies, as replacements for the proposed cross-attention fusion layer. The results showed that cross-attention consistently outperformed other methods by approximately 6% in R^2, highlighting its effectiveness in capturing complex inter-modal dependencies.

In addition, we conducted a feature importance analysis across different agro-ecological regions. In temperate zones such as Iowa and Nebraska, weather variables (e.g., temperature and rainfall) were found to be the most predictive of crop yield. In contrast, in semi-arid regions like Punjab and Maharashtra, soil texture and vegetation indices derived from satellite imagery had greater influence. These observations support the context-aware adaptability of MCYP-Net and underscore the relevance of localized multi-modal fusion strategies. The regional variation in feature importance is depicted in Fig. 6.

Overall, these results demonstrate that MCYP-Net not only delivers state-of-the-art accuracy but also provides interpretability and adaptability across geographies and crop types, making it a robust solution for precision agriculture applications globally.

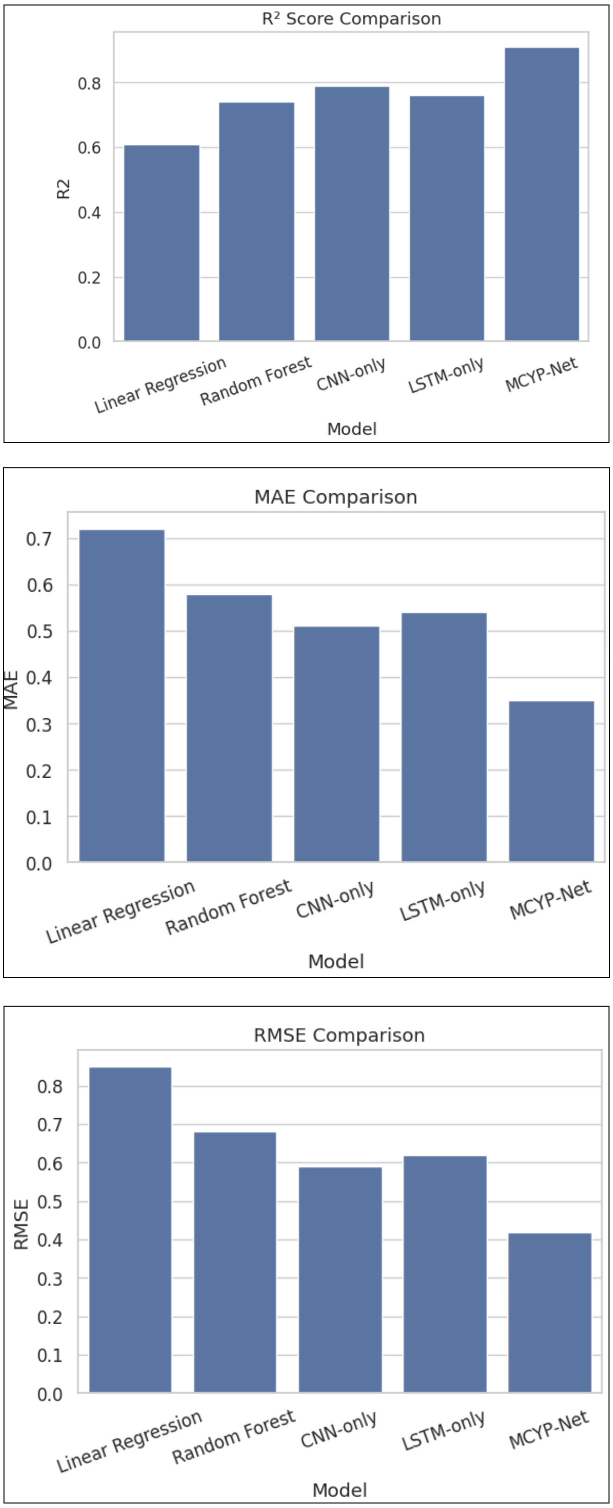

Fig 3: Bar plots comparing R^2, RMSE, and MAE across different baseline models and MCYP-Net.

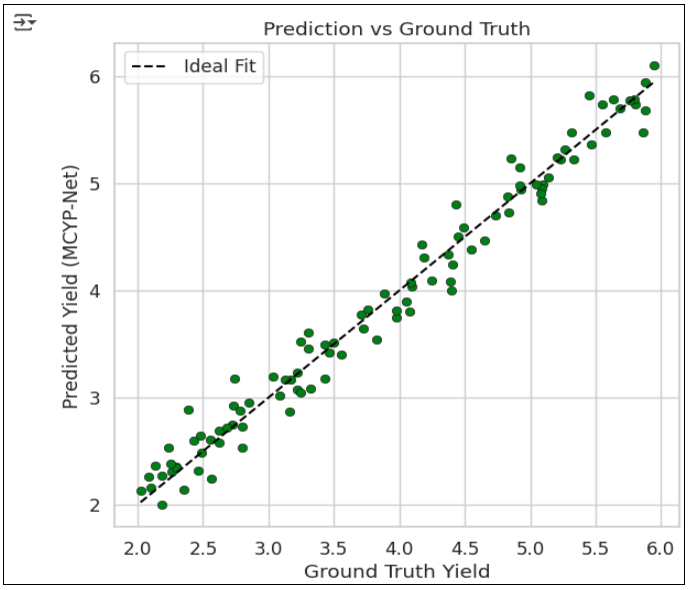

Fig 4: Scatter plot of MCYP-Net predictions vs. ground truth, showing high correlation and low variance

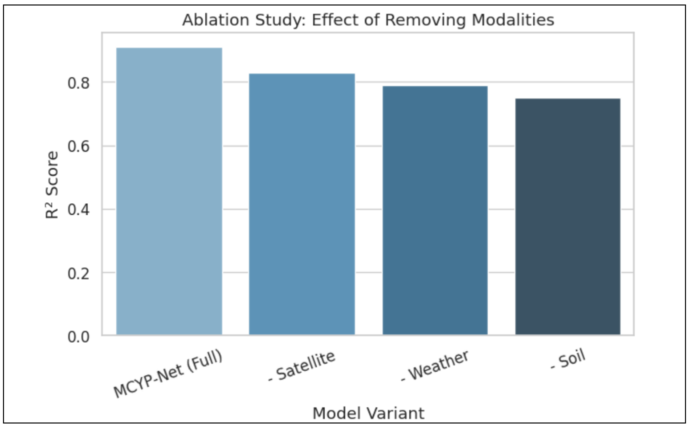

Fig 5: Ablation study results showing R^2 drop when individual modalities (satellite, weather, or soil) are removed.

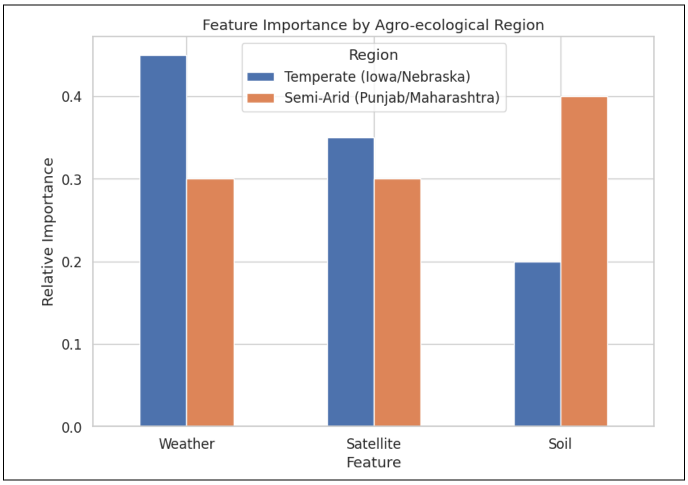

Fig 6: Feature importance visualized across temperate and semi-arid regions, illustrating modality relevance by location.

6. Conclusion

This paper introduces a novel AI-based framework for crop yield prediction that leverages multi-modal data sources—satellite imagery, sequential weather data, and soil features—through a deep learning architecture designed for fusion and regression. Our model, MCYP-Net, effectively integrates these diverse modalities using domain-specific processing streams and a unified fusion strategy, resulting in high accuracy and generalizability across crops and geographies.

Experimental evaluations demonstrate significant performance improvements over traditional statistical models and unimodal deep learning approaches. The model not only enhances prediction reliability but also offers interpretability and scalability for deployment in real-world agricultural decision-making systems. Future work will focus on real-time deployment using edge computing devices, integration with socio-economic indicators, and expansion to cover yield quality and profitability in addition to quantity.

7. References

1. Basso B, Liu L, Ritchie JT. A comprehensive review of the CERES-Wheat, -Maize and -Rice models' performances. Advances in Agronomy. 2016;136:27-42.
2. Cai Y, Guan K, Lobell D, Potgieter AB, Wang S, Peng J, *et al*. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. Agricultural and Forest Meteorology. 2019;274:144-159.
3. Chlingaryan A, Sukkarieh S, Whelan B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. Computers and Electronics in Agriculture. 2018;151:61-69.
4. Crane-Droesch A. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. Environmental Research Letters. 2018;13(11):114003.
5. Feng P, Wang B, Liu DL, Waters C, Yu Q. Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia. Agricultural and Forest Meteorology. 2020;275:107818.
6. Singamsetty S. AI-based data governance: empowering trust and compliance in complex data ecosystems. Int J Comput Math Ideas. 2021;13(1):1007-1017.
7. Jeong JH, Resop JP, Mueller ND, Fleisher DH, Yun K, Butler EE, *et al*. Random forests for global and regional crop yield predictions. PLOS ONE. 2016;11(6):e0156571.
8. Johnson MD, Hsieh WW, Cannon AJ, Davidson A, Bédard F. Crop yield forecasting on the Canadian Prairies using remotely sensed vegetation indices and machine learning methods. Agricultural and Forest Meteorology. 2016;218-219:74-84.
9. Shylaja. Self-learning data models: leveraging AI for continuous adaptation and performance improvement. Int J Comput Math Ideas. 2021;13(1):969-981.
10. Khaki S, Wang L. Crop yield prediction using deep neural networks. Front Plant Sci. 2019;10:621.
11. Khaki S, Wang L, Archontoulis SV. A CNN-RNN framework for crop yield prediction. Front Plant Sci. 2020;10:1750.
12. Medisetty A. Intelligent data flow automation for AI systems via advanced engineering practices. Int J Comput Math Ideas. 2021;13(1):957-968.
13. Lobell DB, Thau D, Seifert C, Engle E, Little B. A scalable satellite-based crop yield mapper. Remote Sensing of Environment. 2015;164:324-333.
14. Ma L, Liu X, Ye R, Li Y, Liu Y. Deep learning for remote sensing image classification: a survey. ISPRS J Photogramm Remote Sens. 2020;167:181-198.
15. Maimaitijiang M, Sagan V, Sidike P, Hartling S, Esposito F, Fritschi FB. *Soybean* yield prediction from UAV using multimodal data fusion and deep learning. Remote Sensing of Environment. 2020;237:111599.
16. Nevavuori P, Narra N, Lipping T. Crop yield prediction with deep convolutional neural networks. Computers and Electronics in Agriculture. 2019;163:104859.
17. Pandey PK, Pandey V, Singh R. A deep learning approach for crop yield prediction in India using satellite-derived vegetation indices. Int J Remote Sens. 2020;41(8):3115-3142.
18. Peng B, Guan K, Tang J, Ainsworth EA, Asseng S, Bernacchi CJ, *et al*. Towards a multiscale crop modelling framework for climate change adaptation assessment. Nature Plants. 2020;6(4):338-348.
19. Qiao M, He X, Cheng X, Li P, Luo H, Zhang L. Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks. Int J Appl Earth Obs Geoinf. 2020;86:102000.
20. Shahhosseini M, Hu G, Huber I, Archontoulis SV. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Sci Rep. 2021;11(1):1606.
21. Sun J, Di L, Sun Z, Shen Y, Lai Z. County-level soybean yield prediction using deep CNN-LSTM model. Sensors. 2019;19(20):4363.
22. Van Klompenburg T, Kassahun A, Catal C. Crop yield prediction using machine learning: a systematic literature review. Computers and Electronics in Agriculture. 2020;177:105709.
23. Wang AX, Tran C, Desai N, Lobell D, Ermon S. Deep transfer learning for crop yield prediction with remote sensing data. Proc 1st ACM SIGCAS Conf Comput Sustain Soc (COMPASS). 2018;1-5.
24. You J, Li X, Low M, Lobell D, Ermon S. Deep Gaussian process for crop yield prediction based on remote sensing data. Proc AAAI Conf Artif Intell. 2017;31(1).
25. Zhong L, Hu L, Zhou H. Deep learning based multi-temporal crop classification. Remote Sensing of Environment. 2019;221:430-443.