A SPATIAL HYBRID MODEL FOR CROP YIELD PREDICTION IN WESTERN AUSTRALIA

Muhammad Ibrahim[1], Balwinder Singh[1], Rodrigo Pires[1], Julia Easton[2], Hasnein Tareque[1]

[1]Department of Primary Industries and Regional Development, 1 Nash Street, Perth, WA, Australia

(muhammad.ibrahim, balwinder.singh, rodrigo.pires, hasnein.tareque) @dpird.wa.gov.au

[2] The Centre for Crop and Disease Management, Curtin University, Bentley, Perth, WA, Australia, Julia.Easton@curtin.edu.au

***Abstract*—Accurate crop yield estimation is crucial for ensuring food security and policy making; however, traditional methods are labor-intensive, and machine learning models often struggle to capture the complex interactions between soil, weather, and temporal data. To address these challenges, this paper first introduces a comprehensive resource—named the WA Paddock Dataset—comprising soil, climate, and satellite data for approximately 450,000 paddocks across Western Australia (WA) over a three-year period. This dataset is designed to enable high-resolution and large-scale modeling and supports other diverse research applications. Building on this resource, we present a novel hybrid regression model for predicting paddock-level crop yield that integrates MAMBA blocks, Transformer attention mechanisms, and Slot Attention to effectively capture spatial and temporal intricacies. This architecture effectively captures spatial and temporal complexities by leveraging diverse geospatial data, including soil properties, weather patterns, and Sentinel-2 imagery, to enhance predictive accuracy. Evaluation against classical machine learning models and ResNet50 demonstrates that our hybrid model significantly improves accuracy while achieving faster inference speeds compared to ResNet50 and some traditional approaches. These results establish the proposed method as a robust and efficient solution for precision agriculture and guiding impactful policy development at regional scale. Our code and relevant dataset will be made publicly available through GitHub repository.**

***Index Terms*—Crop yield estimation, hybrid model, MAMBA, Transformer, slot attention, geospatial, weather, Sentinel-2, ResNet-50, WA Paddock Dataset.**

I. Introduction

Accurate crop yield estimation is crucial for ensuring food security, optimizing agricultural production, and shaping effective policies [1, 2]. Crop yield is influenced by complex interactions between soil properties, weather conditions, and temporal factors, making traditional manual methods labor-intensive, time-consuming, and prone to estimation [3, 4]. Recent advancements in remote sensing and image processing, leveraging data from RGB, multispectral, infrared, and near-infrared satellite imagery, have automated yield estimation with improved accuracy [3, 5, 6]. However, these conventional methods often fail to fully capture the high-dimensional, nonlinear relationships inherent in agricultural data, including critical soil properties (e.g., organic carbon and moisture) and dynamic weather factors (e.g., temperature, rainfall, and humidity).

This study received funding from Food Agility CRC Australia.

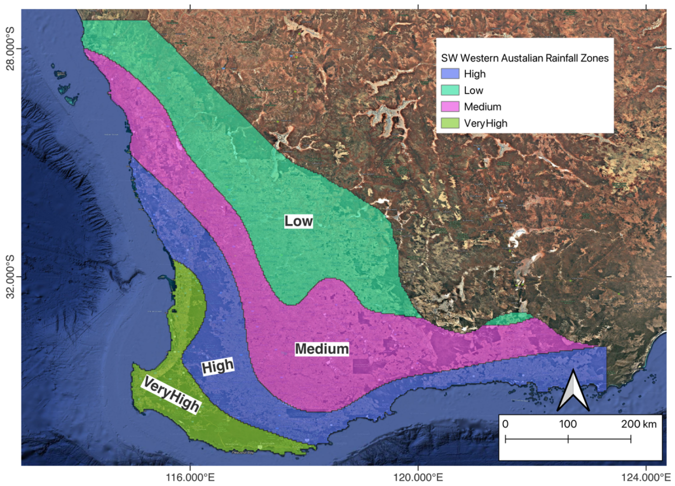

Fig. 1. Map of Western Australia delineating rainfall zones categorized as Very High, High, Medium, and Low [7]. These zones represent varying precipitation levels, providing a framework for analyzing crop yield variability. The basemap utilized is sourced from Google Earth [8].

The evolution of crop yield prediction has seen a shift from traditional methods to classical machine learning models [9], such as Random Forest Regression [10], Support Vector Regression (SVR) [11], and ensemble learning approaches [12], which have improved predictive performance by leveraging structured data and advanced algorithms [13–15]. Nonetheless, their inability to effectively capture complex spatial, temporal, and cross-domain interactions has led to a growing interest in leveraging deep learning methodologies. These methods, such as Multi-Layer Perceptrons (MLP) and Convolutional Neural Networks (CNN), excel at learning complex and intricate data patterns, offering predictive capabilities for large-scale crop yield prediction and environmental monitoring [2, 16]. Despite their success, cutting-edge architectures like Transformers [17] and MAMBA models [18] remain underexplored in agriculture. These advanced models, widely used in natural language processing (NLP) and computer vision, have yet to be evaluated for their ability to model the complex interplay of agricultural factors. Furthermore, the lack of large-scale, paddock-level crop yield datasets limits the application of advanced models and impedes prediction accuracy [21, 22]. This gap underscores the urgent need for innovative methodologies and comprehensive datasets to advance precision agriculture. Addressing this gap, our first contribution is the creation

TABLE I
SUMMARY OF WA PADDOCK DATASET STATISTICS FROM 2020 TO 2022. THE UNIQUE CROP NAMES ARE REPRESENTED USING SHORT FORMS: W (WHEAT), B (BARLEY), AND C (CANOLA).

Year	Avg. Yield (t/ha)	Max Yield (t/ha)	Total Area (ha)	# Paddocks	Total Prod. (t)	Crop Names	# Weather Features	# Soil Features	# Sentinel-2 Bands Features
2022	3.074	7.95	8,127,189.4	151,265	24,622,366	W, B, C	18	119	5
2021	2.789	6.4	8,036,881.4	147,681	22,165,259	W, B, C,	18	119	5
2020	2.085	6.06	7,629,099.8	138,517	15,281,462	W, B, C	18	119	5

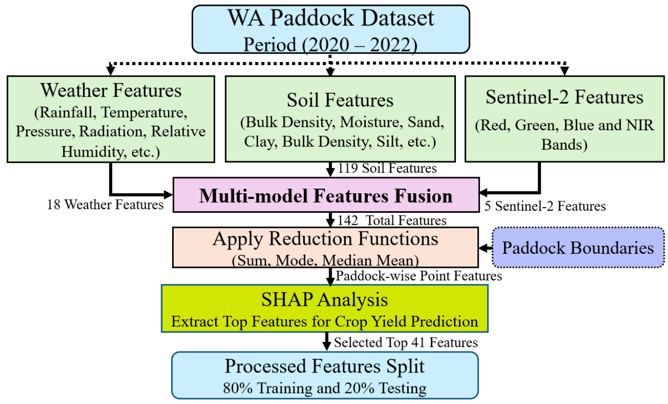

Fig. 2. Overview of the WA Paddock Dataset and Feature Processing Pipeline. The dataset integrates weather, soil, and Sentinel-2 features [19, 20].

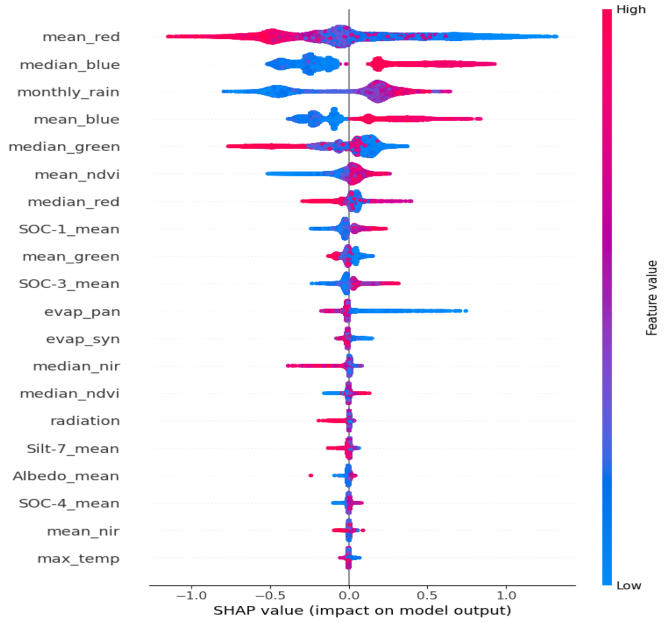

Fig. 3. SHAP analysis summary plot for the WA Paddock Dataset, highlighting the impact of key features on crop yield prediction. Showing the top 20 out of 41 selected features for improved visualization. SOC: Soil Organic Carbon and NIR: Near Infra-Red Band.

of the WA Paddock Dataset, comprising data for 450,000 paddocks across Western Australia from 2020–2022. This dataset integrates soil properties [23], weather data [24], and Sentinel-2 satellite imagery [19, 20] in raw and processed formats. Processed data aggregates key attributes (e.g., soil and climate metrics) using reduction functions such as SUM, MEDIAN, and MEAN, ensuring flexibility for diverse research needs. Additionally, SHAP analysis [25] identifies the most relevant features for crop yield prediction, optimizing model performance. The dataset structure is illustrated in Fig. 2.

Our second major contribution is the development of a novel hybrid regression model for paddock-level crop yield prediction, integrating MAMBA streams, Transformer attention mechanisms, and Slot Attention [26]. This architecture combines coarse- and fine-grained learning, leveraging MAMBA's robust modeling capabilities [18] and Transformers' attention mechanisms to effectively process spatial and temporal complexities. Slot Attention [26] enhances feature representation by dynamically prioritizing critical patterns, delivering a comprehensive and accurate approach to yield prediction.

We evaluate the performance of the hybrid model against classical methods (e.g., K-Nearest Neighbors [27], Gradient Boosting [28], Support Vector Regression (SVR) [29], and Random Forest) [10] and the deep learning benchmark ResNet50 [30] using statistical parameters like R^2, Mean Squared Error (MSE), and inference time. The proposed model consistently surpasses these baselines in accuracy and achieves faster inference speeds than ResNet50 [30] and classical machine learning models like SVR [29], highlighting its robustness and scalable implementation potential for real-world agricultural applications.

II. DATASET DESCRIPTION

A. Study Site

This study focuses on the WA Rainfall based agriculture region [7], a key hub for agricultural production. The region primarily cultivates Barley, Canola, and Wheat crops, with the growing season aligned with the available rainfall. It is divided into four rainfall zones based on annual precipitation levels, as shown in Fig. 1: Very High, High, Medium, and Low. While the Very High Rainfall Zone includes urban areas like Perth with limited agricultural activity, the other zones are predominantly utilized for farming. Crop yield in this region is shaped by complex interactions between weather and soil properties. To analyze these factors, we developed the WA Paddock Dataset, integrating diverse environmental and agricultural data. This dataset serves as a robust foundation for yield prediction and model validation, leveraging the region's climatic and soil variability.

B. WA Paddock Dataset

The WA Paddock Dataset, covering years 2020-2022, is a comprehensive resource developed for advanced crop yield modeling. This dataset integrates weather [24], soil [23], and satellite data [19, 20], capturing 142 key features. The dataset is summarized in Table I.

The raw dataset includes 18 weather-related features [24], such as total rainfall, minimum and maximum temperature, evapotranspiration (syn, pan, Morton lake), radiation, relative humidity (at Tmax and Tmin), mean sea level pressure (MSLP), vapor pressure, and Morton indicators (actual, potential, wet). In terms of soil attributes [23], 119 features are captured for bulk density, organic carbon, organic carbon fraction, moisture, sand, silt, clay, soil albedo [31][20], soil depth,

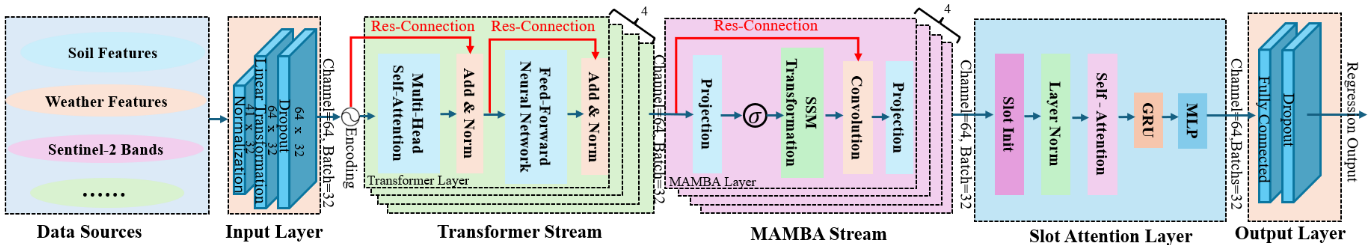

Fig. 4. Overview of the proposed architecture, highlighting the five main components: Input Layer, Transformer Stream, MAMBA Stream, Slot Attention Block, and Regression Output Layer.

coarse fragments, drained upper limit, and other properties such as Atlas, Illite, Kaolinite, and Smectite. Additionally, SMOS data (maximum, mean, median, minimum) is extracted, along with soil color [32], soil classification, digital elevation surface (DES), and digital elevation models (DEMs) [33]. From Sentinel-2 [19] satellite imagery, five key features are derived, including spectral bands and NDVI values.

The raw data are transformed into paddock-wise point features based on WA paddock boundaries [34][35], as outlined in Fig. 2. Reduction functions such as SUM, MODE, MEDIAN, and MEAN summarize the features within each paddock. The fused dataset is further refined through SHAP analysis [25] to identify the most critical variables for crop yield prediction. Fig 3 highlights the top features for crop yield prediction identified through SHAP analysis. This dataset provides detailed insights into crop variabilities and yield dynamics, serving as a valuable resource for training and evaluating machine learning models to advance precision agriculture and sustainable farming.

III. Methodology

In this section, we present the methodology of the proposed hybrid architecture, which integrates multiple advanced neural components to predict crop yield based on multimodal data. The methodology is structured into five components: Input Layer, Transformer Stream, MAMBA Stream, Slot Attention Block, and Regression Output Layer. Each component is designed to process and refine input features, enabling robust and accurate crop predictions at the paddock level. The overall architecture is depicted in Fig. 4.

A. Input Layer

The Input Layer is designed to preprocess input data, which includes Sentinel-2 features, soil characteristics, and weather data. The input features are normalized and mapped to 64 feature channels through a linear transformation layer. The transformed features can be represented as $Z \in \mathbb{R}^{B \times d}$, where B denotes the batch size.

B. Transformer Stream

The Transformer stream is responsible for capturing global dependencies and contextual relationships within the data. It is composed of four layers, each containing three blocks of Multi-Head Attention mechanisms and Feed-Forward Neural Networks [17], as illustrated in Fig. 5. The Multi-Head Attention mechanism is defined as:

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V \tag{1}$$

Where $Q = ZW_Q$, $K = ZW_K$, $V = ZW_V$, Z is the input data, and W represents the learnable weights. This

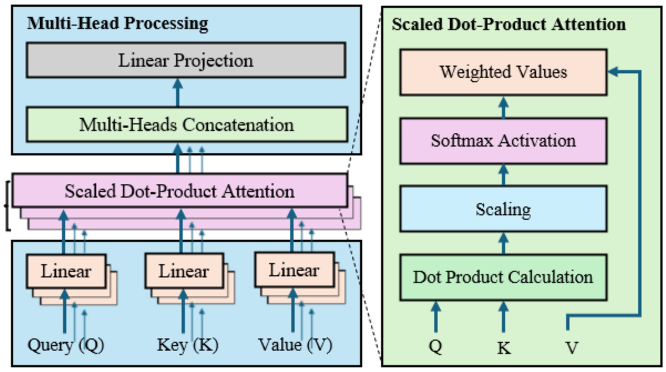

Fig. 5. Detailed view of the Multi-Head Attention mechanism used in the Transformer Block [17].

mechanism enables the model to focus on different parts of the input sequence when deriving contextual relationships. Residual connections and layer normalization are applied after each sub-layer to ensure stable training and performance.

C. MAMBA Stream

Similar to the previous stream, this stream comprises four layers, each constructed with three MAMBA blocks. The MAMBA Block enhances temporal and spatial feature extraction by combining Sequence Transformation and Convolutional operations [18]. The input features are processed via parallel branches, enabling the efficient extraction of complex details. This design makes MAMBA ideally suited for tasks requiring comprehensive feature understanding and robust modeling. The Sequence-to-Sequence Mapping (SSM) in the MAMBA Block is defined as:

$$X_{\text{SSM}} = \sigma(W_s X + b_s) \tag{2}$$

where X represents the input features, W_s and b_s are learnable parameters, and σ is a non-linear activation function such as GELU. The outputs from the SSM and linear branches are combined and passed through a convolutional layer to enhance feature expressiveness. Residual connections are employed to ensure gradient flow [18].

D. Slot Attention Block

The Slot Attention Block dynamically routes features into fixed-size slots, enabling a structured representation of input data [26]. The iterative attention mechanism refines slot representations over multiple steps. Slots are initialized as learnable parameters and iteratively updated using attention weights computed between the input features and the current slot states. Each slot aggregates relevant input information and is further refined using a GRU and an MLP. This mechanism ensures a compact and interpretable representation of the input features.

TABLE II

PERFORMANCE COMPARISON OF YIELD PREDICTION METHODS ON THE WA PADDOCK DATASET. THE COMPARISON METRICS INCLUDE R^2 ACCURACY, MEAN SQUARED ERROR (MSE), AND INFERENCE TIME. ALL METHODS WERE TRAINED AND TESTED ON THE PROPOSED DATASET.

Method	**Dataset**	**Type**	R^2 **Accuracy (%)**	**MSE**	**Inference Time (µs)**
Linear Regression [36]	WA Rainfall Paddocks	Classical ML	66.6	0.4255	0.06
Ridge Regression [37]	WA Rainfall Paddocks	Classical ML	66.60	0.4255	0.06
Lasso Regression [38]	WA Rainfall Paddocks	Classical ML	30.84	0.8810	0.05
ElasticNet [39]	WA Rainfall Paddocks	Classical ML	41.35	0.7471	**0.049**
Decision Tree [40]	WA Rainfall Paddocks	Classical ML	69.57	0.3877	0.54
K-Nearest Neighbors [27]	WA Rainfall Paddocks	Classical ML	78.21	0.2775	104.33
SVR [29]	WA Rainfall Paddocks	Classical ML	81.69	0.23	7,665.93
Gradient Boosting [28]	WA Rainfall Paddocks	Classical ML	79.11	0.2662	1.15
Random Forest [10]	WA Rainfall Paddocks	Classical ML	84.98	0.1913	33.32
ResNet50 [30]	WA Rainfall Paddocks	Deep Learning	84.21	0.2104	65.10
Proposed Hybrid Base Model	WA Rainfall Paddocks	Deep Learning	85.12	0.1892	27.54
Proposed Hybrid Model	WA Rainfall Paddocks	Deep Learning	**86.43**	**0.1788**	58.69

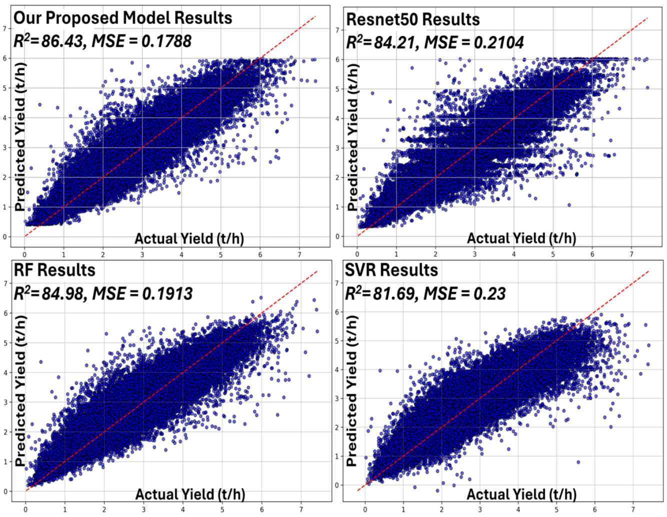

Fig. 6. Visualization of actual vs predicted crop yield for WA Paddock Dataset. The proposed model achieves the best performance with $R^2 = 86.43$ and $MSE = 0.1788$, outperforming ResNet50, RF, and SVR.

E. Regression Output Layer

The Regression Output Layer maps the refined features to a final prediction. The slots S_{final} are pooled into a single representation, $\hat{y} = \text{ReLU}(\text{MeanPooling}(S_{\text{final}})W + b)$, where $\hat{y} \in \mathbb{R}$ represents the predicted crop yield. Dropout is applied to mitigate overfitting.

IV. RESULTS AND DISCUSSION

This section evaluates the proposed hybrid model against state-of-the-art classical and deep learning methods using the WA Paddock Dataset, employing R^2 accuracy, MSE, and inference time as evaluation metrics. Fig. 6 visually compares the top four models: the hybrid model, Random Forest [10], ResNet50 [30], and SVR [29], highlighting the hybrid model's better alignment with ground truth. Table II summarizes the results, showing the hybrid model achieves the highest R^2 (86.43%) and lowest MSE (0.1788). Random Forest [10] follows with an R^2 of 84.98% and an MSE of 0.1913, while ResNet50 [30] and SVR lag at R^2 values of 84.21% and 81.69%, respectively. The hybrid model also demonstrates competitive inference time, making it suitable for real-time precision agriculture applications.

A streamlined variant of the hybrid model, with reduced MAMBA and Transformer layers, achieves faster inference with only a 1.3% decrease in R^2 accuracy, demonstrating scalability for computationally constrained scenarios. Additionally, SHAP analysis [25] highlights the most critical features derived from weather [24], soil [23], and Sentinel-2 data [19, 20], enhancing the interpretability of the model and establishing the WA Paddock Dataset as a reliable resource for precision agriculture research. The proposed hybrid model achieves an optimal balance of accuracy, efficiency, and scalability. Paired with the comprehensive dataset, it provides a practical and robust solution to predict crop yield.

V. CONCLUSION

This paper presents a validated hybrid regression model combining MAMBA blocks, Transformer attention mechanisms, and Slot Attention to process diverse agricultural data. By effectively modeling spatial and temporal relationships, the model achieves improvements in accuracy and inference speed, outperforming classical models and deep learning benchmarks like ResNet50. The introduction of the WA Paddock Dataset further strengthens this research, providing a novel resource for high-resolution and large-scale agricultural modeling. The integration of SHAP analysis refines the model by identifying the most critical features for yield prediction. Overall, this study highlights the transformative potential of integrating geospatial and geophysical data with advanced machine learning and deep learning models. These methodologies offer precise, efficient, and scalable solutions to address the pressing challenges of global food security and sustainable agricultural production.

Acknowledgment: This study was funded by Food Agility Cooperative Research Centre (CRC) Australia. The CRC Program supports industry-led collaborations between industry, researchers and the community. We also acknowledge the contributions of the Department of Primary Industries and Regional Development (DPIRD), WA, for their support in providing technical expertise, access to resources, and facilitation of research activities.

REFERENCES

[1] Q. Xin, P. Gong, C. Yu, L. Yu, M. Broich, A. E. Suyker, and R. B. Myneni, "A production efficiency model-based method for satellite estimates of corn and soybean yields in the midwestern us," *Remote Sensing*, vol. 5, no. 11, pp. 5926–5943, 2013.

[2] S. Sarkar, A. Dey, R. Pradhan, U. M. Sarkar, C. Chatterjee, A. Mondal, and P. Mitra, "Crop yield prediction using multimodal meta-transformer and temporal graph neural networks," *IEEE Transactions on AgriFood Electronics*, 2024.

[3] N. N. Bhookya, R. Malmathanraj, and P. Palanisamy, “Yield estimation of chilli crop using image processing techniques,” in *2020 6th international conference on advanced computing and communication systems (ICACCS)*. IEEE, 2020, pp. 200–204.

[4] N. Goel, S. Kaur, and Y. Kumar, “Machine learning-based remote monitoring and predictive analytics system for crop and livestock,” in *AI, Edge and IoT-based Smart Agriculture*. Elsevier, 2022, pp. 395–407.

[5] X. Zhang, A. Toudeshki, R. Ehsani, H. Li, W. Zhang, and R. Ma, “Yield estimation of citrus fruit using rapid image processing in natural background,” *Smart Agricultural Technology*, vol. 2, p. 100027, 2022.

[6] A. Gong, J. Yu, Y. He, and Z. Qiu, “Citrus yield estimation based on images processed by an android mobile phone,” *Biosystems Engineering*, vol. 115, no. 2, pp. 162–170, 2013.

[7] D. Geographic Information Services, “Crop variety testing (cvt) zones of western australia,” 2016, map. [Online]. Available: https://library.dpird.wa.gov.au/gis_maps/1

[8] Google Earth, “Google earth,” https://earth.google.com/intl/earth, 2025, accessed: 2025-01-09.

[9] E. Elbasi, C. Zaki, A. E. Topcu, W. Abdelbaki, A. I. Zreikat, E. Cina, A. Shdefat, and L. Saker, “Crop prediction model using machine learning algorithms,” *Applied Sciences*, vol. 13, no. 16, p. 9288, 2023.

[10] L. Breiman, “Random forests,” *Machine learning*, vol. 45, pp. 5–32, 2001.

[11] N. Jayasuriya, Y. Guo, W. Hu, and O. Ghannoum, “Image based crop monitoring technologies in protected horticulture: A review,” *arXiv preprint arXiv:2401.13928*, 2024.

[12] M. Majnoun Hosseini, M. J. Valadan Zoej, A. Taheri Dehkordi, and E. Ghaderpour, “Cropping intensity mapping in sentinel-2 and landsat-8/9 remote sensing data using temporal transfer of a stacked ensemble machine learning model within google earth engine,” *Geocarto International*, vol. 39, no. 1, p. 2387786, 2024.

[13] D. Pathak, M. Miranda, F. Mena, C. Sanchez, P. Helber, B. Bischke, P. Habelitz, H. Najjar, J. Siddamsetty, D. Arenas *et al.*, “Predicting crop yield with machine learning: An extensive analysis of input modalities and models on a field and sub-field level,” in *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2023, pp. 2767–2770.

[14] Y. Zhang and B. Wallace, “A sensitivity analysis of (and practitioners’ guide to) convolutional neural networks for sentence classification,” *arXiv preprint arXiv:1510.03820*, 2015.

[15] P. Helber, B. Bischke, C. Packbier, P. Habelitz, and F. Seefeldt, “An operational approach to large-scale crop yield prediction with spatio-temporal machine learning models,” in *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2024, pp. 4299–4302.

[16] D. Tedesco-Oliveira, R. P. da Silva, W. M. Jr, and C. Zerbato, “Convolutional neural networks in predicting cotton yield from images of commercial fields,” *Computers and Electronics in Agriculture*, vol. 171, p. 105307, 2020.

[17] A. Vaswani, “Attention is all you need,” *Advances in Neural Information Processing Systems*, 2017.

[18] A. Gu and T. Dao, “Mamba: Linear-time sequence modeling with selective state spaces,” *arXiv preprint arXiv:2312.00752*, 2023.

[19] M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort *et al.*, “Sentinel-2: Esa’s optical high-resolution mission for gmes operational services,” *Remote sensing of Environment*, vol. 120, pp. 25–36, 2012.

[20] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, “Google earth engine: Planetary-scale geospatial analysis for everyone,” *Remote sensing of Environment*, vol. 202, pp. 18–27, 2017.

[21] W. Xu, P. Chen, Y. Zhan, S. Chen, L. Zhang, and Y. Lan, “Cotton yield estimation model based on machine learning using time series uav remote sensing data,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 104, p. 102511, 2021.

[22] K. Meghraoui, I. Sebari, J. Pilz, K. Ait El Kadi, and S. Bensiali, “Applied deep learning-based crop yield prediction: A systematic analysis of current developments and potential challenges,” *Technologies*, vol. 12, no. 4, 2024. [Online]. Available: https://www.mdpi.com/2227-7080/12/4/43

[23] B. P. Malone, R. D. Searle, M. Stenson, D. McJannet, P. Zund, M. R. Dobarco, A. M.-C. Wadoux, B. Minasny, A. McBratney, and M. Grundy, “Update and expansion of the soil and landscape grid of australia,” *Available at SSRN 4987535*.

[24] S. J. Jeffrey, J. O. Carter, K. B. Moodie, and A. R. Beswick, “Using spatial interpolation to construct a comprehensive archive of australian climate data,” *Environmental Modelling & Software*, vol. 16, no. 4, pp. 309–330, 2001.

[25] S. Lundberg, “A unified approach to interpreting model predictions,” *arXiv preprint arXiv:1705.07874*, 2017.

[26] F. Locatello, D. Weissenborn, T. Unterthiner, A. Mahendran, G. Heigold, J. Uszkoreit, A. Dosovitskiy, and T. Kipf, “Object-centric learning with slot attention,” *Advances in neural information processing systems*, vol. 33, pp. 11 525–11 538, 2020.

[27] Y. Song, J. Liang, J. Lu, and X. Zhao, “An efficient instance selection algorithm for k nearest neighbor regression,” *Neurocomputing*, vol. 251, pp. 26–34, 2017.

[28] A. Natekin and A. Knoll, “Gradient boosting machines, a tutorial,” *Frontiers in neurorobotics*, vol. 7, p. 21, 2013.

[29] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, “Support vector regression machines,” *Advances in neural information processing systems*, vol. 9, 1996.

[30] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[31] C. Schaaf and Z. Wang, “Mcd43a3 modis/terra+aqua brdf/albedo daily l3 global - 500m v006,” https://doi.org/10.5067/MODIS/MCD43A3.006, 2015, accessed 2025-01-10.

[32] AusSoilsDSM, “Soil colour,” https://aussoilsdsm.esoil.io/slga-version-2-products/soil-colour, 2025, accessed: 2025-01-09.

[33] J. Gallant, N. Wilson, T. Dowling, A. Read, and C. Inskeep, “Srtm-derived 1 second digital elevation models version 1.0,” Geoscience Australia, Canberra, Tech. Rep. Record 1, 2011. [Online]. Available: https://pid.geoscience.gov.au/dataset/ga/72759

[34] F. Waldner and F. I. Diakogiannis, “Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network,” *Remote sensing of environment*, vol. 245, p. 111741, 2020.

[35] Digital Agriculture Services, “Grain intelligence,” n.d., accessed: 2025-01-09. [Online]. Available: https://digitalagricultureservices.com/products/grain-intelligence

[36] O. O. Aalen, “A linear regression model for the analysis of life times,” *Statistics in medicine*, vol. 8, no. 8, pp. 907–925, 1989.

[37] G. C. McDonald, “Ridge regression,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 1, no. 1, pp. 93–100, 2009.

[38] J. Ranstam and J. A. Cook, “Lasso regression,” *Journal of British Surgery*, vol. 105, no. 10, pp. 1348–1348, 2018.

[39] C. Hans, “Elastic net regression modeling with the orthant normal prior,” *Journal of the American Statistical Association*, vol. 106, no. 496, pp. 1383–1393, 2011.

[40] W.-Y. Loh, “Classification and regression trees,” *Wiley interdisciplinary reviews: data mining and knowledge discovery*, vol. 1, no. 1, pp. 14–23, 2011.