Soft Computing Fusion with Applications

www.scfa.reapress.com

Soft. Comput. Fusion. Appl. Vol. 1, No. 4 (2024) 253–262.

Paper Type: Original Article

Predictive Modeling of Agricultural Yield Using Multi-Source Geospatial Data

Nomaan Siraj[1], Swayam Swaroop[1,*], Tanya Raj[1], Anubhav Mishra[1]

[1] Department of Computer Sciencekalinga, Kalinga Institute of Industrial Technology (KIIT) University, Bhubaneswar, Odisha, India; 21052085@kiit.ac.in; 2105076@kiit.ac.in; 21052888@kiit.ac.in; 2105013@kiit.ac.in.

Citation:

Received: 28 June 2024 *Revised: 16 September 2024* *Accepted: 13 November 2024*	Siraj, N., Swaroop, S., Raj, T., & Mishra, A. (2024). Predictive modeling of agricultural yield using multi-source geospatial data. *Soft computing fusion with applications, 1*(4), 253-262.

Abstract

Accurate prediction of agricultural yield is crucial for ensuring food security and optimizing resource allocation. This project aims to develop a robust predictive model that leverages the power of remote sensing, weather data, and soil information to estimate crop yield accurately. By integrating advanced machine learning and deep learning techniques with geospatial analysis, we strive to improve the precision and reliability of yield forecasts. The proposed methodology involves several key steps: 1) data acquisition and preprocessing, 2) model development and training, and 3) deployment and visualization.

Keywords: Agricultural yield prediction, Remote sensing, Machine learning, Deep learning, Geospatial analysis, Satellite imagery, Predictive modeling.

1 | Introduction

The precision and accessibility of earth observations have steadily improved due to ongoing advancements in space-based remote sensing, like the launch of the Planet and open-access Sentinel satellite constellations [1]–[3]. Notably, sa]tellites with high revisit frequencies provide valuable insights into phenomena with complex temporal dynamics. Crop mapping, which is the main focus of this paper, leverages these temporal patterns and holds significant financial and environmental implications. Remote monitoring of agricultural lands is essential for fair distribution of agricultural subsidies amounting to 50 billion euros annually in Europe and 22 billion euros in the US and for ensuring adherence to optimal crop rotation practices [4]–[6].

In a broader context, automated Satellite Image Time Series (SITS) analysis has significant value across diverse applications, such as tracking urban expansion and monitoring deforestation [7], [8]. The content and boundaries of agricultural parcels can be observed as panoptic segmentation of an image sequence. Panoptic segmentation involves assigning each pixel a specific class and unique instance label, making it a well-

Corresponding Author: 2105076@kiit.ac.in
https://doi.org/10.22105/scfa.v1i4.71

established task in computer vision [9]. However, this task differs fundamentally from natural image or video sequences for SITS. Unlike videos, which require tracking objects over time and space; SITS involve static targets within a geo-referenced frame, eliminating the need for spatial tracking. Additionally, SITS operate on a consistent temporal frame, where acquisition timing provides critical data for modeling temporal dynamics, unlike video frames, where sequence numbers are often arbitrary. Finally, while Earth surface objects rarely obscure each other, analysis of SITS can be complicated by varying cloud cover.

This paper addresses the challenge of distinguishing individual agricultural parcels by learning intricate and specific temporal, spatial, and spectral patterns unique to agricultural monitoring, such as plant growth stages, subtle border details, and rapid human activities like harvesting [10], [11]. Although deep neural networks have proven effective for pixel-level classification of such complex patterns, no approach is currently tailored for detecting individual objects in SITS. Previous methods in instance segmentation have focused on single satellite images only. Generally, specialized remote sensing techniques are limited to semantic or single-image instance segmentation, while panoptic segmentation models from computer vision require substantial adaptation to handle SITS [12]–[14].

We introduce U-net with Temporal Attention Encoder (U-TAE) to tackle this. This innovative spatio-temporal encoder integrates multi-scale spatial convolutions and a temporal self-attention mechanism, allowing it to prioritize the most significant time acquisitions [15]. Unlike convolutional-recurrent methods that only capture temporal features at high or low spatial resolutions, U-TAE leverages predicted temporal masks to adaptively learn spatiotemporal features across multiple resolutions simultaneously. Additionally, we propose Parcels-as-Points (PaPs), the first end-to-end deep learning solution for panoptic segmentation in SITS, building on the efficient CenterMask network, which we modified for this task [16]–[18]. Finally, we present PASTIS, the first publicly available dataset designed for training and evaluating panoptic segmentation models on SITS, with over 2 billion annotated pixels spanning over 4,000 km². Our approach, tested on PASTIS, outperforms all re-implemented competing methods for semantic segmentation, establishing a new benchmark for SITS panoptic segmentation.

2 | Literature Review

Extensive research exists on encoding satellite image sequences and panoptic segmentation of videos and single satellite images.

Encoding satellite image sequences

Early automated SITS analysis tools used traditional machine learning techniques [19]. However, deep convolutional networks have since enabled the extraction of more detailed spatial features [20]. Initially, temporal aspects were addressed through handcrafted temporal descriptors or probabilistic models, which have been effectively superseded by architectures such as recurrent, convolutional, or differential networks. More recently, attention-based methods have also been introduced. It has been adapted to encode sequences of remote-sensing images and has led to significant progress in pixel-wise and parcel-wise classification. In parallel, hybrid architectures relying on U-Net-type architectures for encoding the spatial dimension and recurrent networks for the temporal dimension are well suited for the semantic segmentation of SITS. In this paper, we propose to combine this hybrid architecture with the promising temporal attention mechanism.

Instance segmentation of satellite images

The initial step in panoptic segmentation is identifying individual instances, known as instance segmentation. Most instance segmentation methods in remote sensing focus on single-image acquisitions [21]. For example, several techniques have been developed to detect individual instances of trees, buildings, or fields. Some approaches begin with a border detection step and require postprocessing to define distinct cases. Other methods use segmentation as a preliminary step to compute cluster-based features, though they do not provide clear mappings of clusters to specific objects. Petitjean and Dacher [22] introduced a segmentation-aided classification method for image time series; however, their approach segments each image separately without consistently identifying objects across the sequence. This paper presents the first end-to-end framework designed to perform semantic and instance segmentation on SITS jointly [23].

Panoptic segmentation of videos

In extensive research on instance segmentation, Mask-RCNN is a leading method for natural images. Recently, Wang et al. [24] introduced CenterMask, a lighter and more efficient single-stage approach, which serves as a foundation for our work in this paper. Several methods aim to extend instance or panoptic segmentation techniques from images to video. However, as discussed in the introduction, SITS differ significantly from natural videos, necessitating specialized algorithmic and architectural modifications.

3 | Figures and Tables

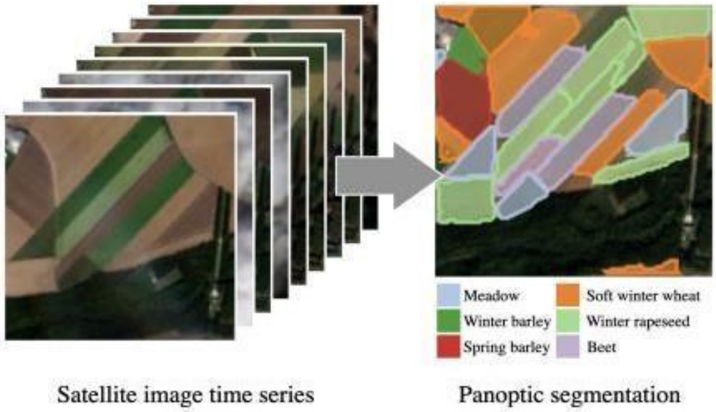

Fig. 1. Schematic diagram of the proposed U-net with temporal attention encoder model architecture for processing satellite image time series, showing the spatial encoder, temporal attention mechanism, and decoder components.

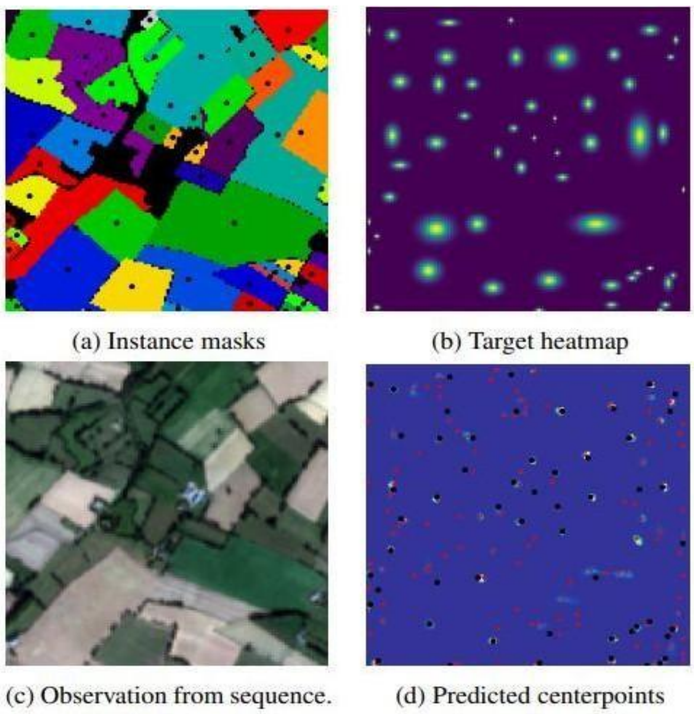

Fig. 2. Panoptic segmentation.

Table 1. Dataset description.

Dataset Listing of Attributes	
Crop	String, crop name
Vriety	String, crop subsidary name
State	String, crops cultivatons/production place
Quantity	Integer, no of qointals, hectars
Production	Integer, no of years production
Season	Date time, medium (No of dayes), long (No of dayes)
Unit	String, tons
Cost	Integer, cost of cutivation and production

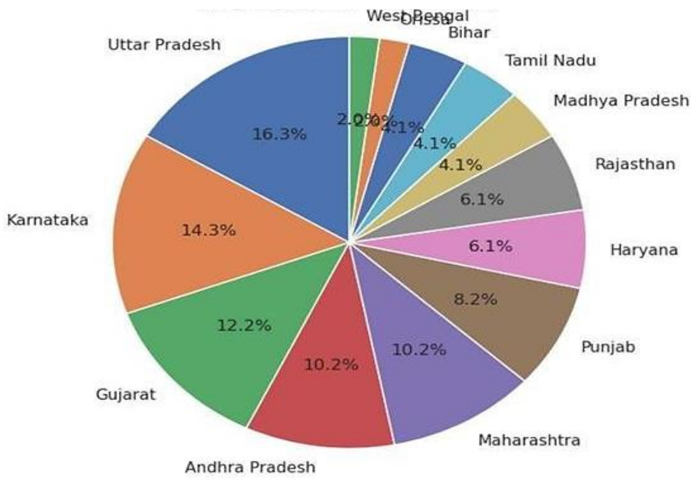

Fig. 3. Distribution of states in the data.

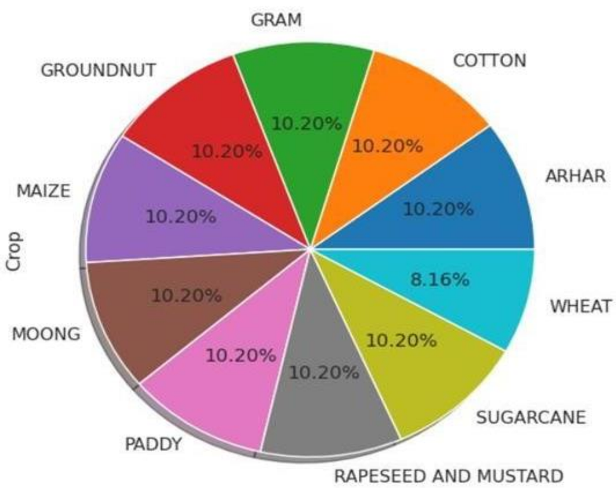

Fig. 4. Top 10 crop in the data.

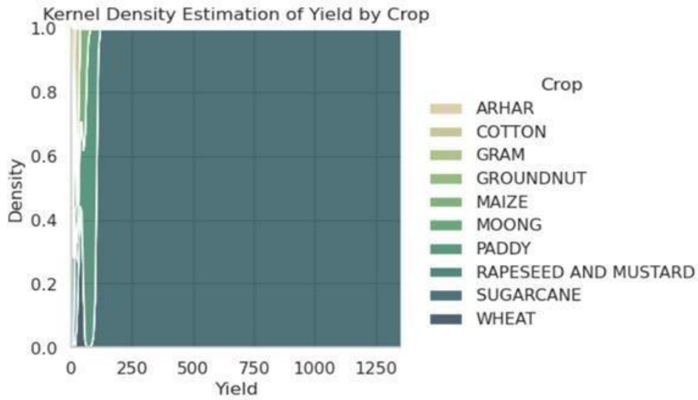

Fig. 5. Kernel density estimation of yield by crop.

Here is a tabular description of the PASTIS dataset as presented in the document:

Table 2. PASTIS dataset that presented in the document.

Attribute	Description
Dataset name	PASTIS
Image resolution	10 meters per pixel
Image shape	Each image sequence: 10 × 128 × 128
Temporal range	Images taken between September 2018 and November 2019
Temporal observation ns	38 to 61 observations per sequence
Spatial coverage	Over 4,000 km² covering regions in
Number of sequences	France 2,433 sequences
Total pixels	Over 2 billion annotated pixels
Spectral channels	10 channels (Non-atmospheric spectral bands of Sentinel-2, after atmospheric correction)
Annotations	Each pixel has a semantic label (18 crop types + background) and a unique instance label for agricultural parcel boundaries
Unique parcels	124,422 parcels with bounding boxes, pixel-precise masks, and crop-type annotations
Cloud cover filtering	Acquisitions with high cloud cover are filtered automatically
Split folds	5-fold cross-validation with 1 km buffer between images in different folds to prevent contamination

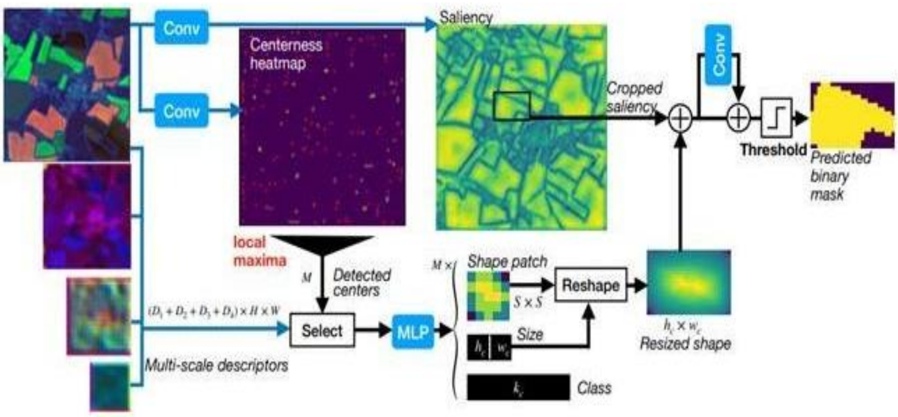

Fig. 6. Performance comparison between U-net with temporal attention encoder and baseline methods on agricultural yield prediction metrics.

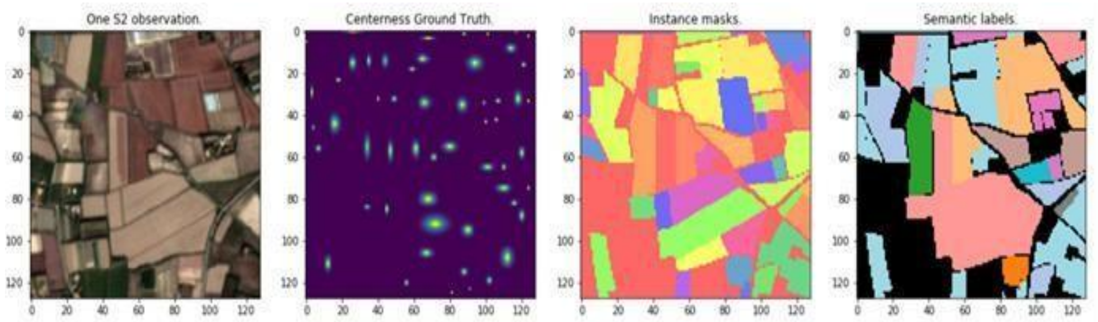

Fig. 7. Qualitative results of panoptic segmentation on satellite imagery.

4 | Variables and Equations

We consider an image time sequence X, organized into a four-dimensional tensor of shape $T \times C \times H \times W$, with T the length of the sequence, C the number of channels, and $H \times W$ spatial extent.

Our model, called U-TAE, encodes a sequence X in three main steps: 1) each image in the sequence is embedded simultaneously and independently using a shared multi-level spatial convolutional encoder, 2) a temporal attention encoder compresses the temporal dimension of the resulting sequence of feature maps into a single feature map per level, and 3) a spatial convolutional decoder generates a final feature map that matches the resolution of the input images.

Spatial encoding

We consider a convolutional en-coder ε with L levels 1, ..., L. Each level is composed of a sequence of convolutions, Rectified Linear Unit (ReLu) activations, and normalizations. Except for the first level, each block starts with a strided convolution, dividing the resolution of the feature maps by a factor 2.

For each time stamp t simultaneously, the encoder ε_t at level / takes as input the feature map of the previous level e_t^{t-1}, and outputs a feature map e_t^t, of size $C \times H \times Wi$ with $H1 = H/2 - 1$ and $W = W/2 - 1$. The resulting feature maps are then temporally stacked into a feature map sequence e^t of size $Tx\ Cix\ H \times W$:

$$e^t = [e_t(e_t^{t-1})]^T{}_{t=0} \ \ t \in [1, L]. \quad (1)$$

With $e^0 = X$ and [0] the concatenation operator along the temporal dimension. When constituting batches, we flatten the temporal and batch dimensions. Since each sequence comprises images acquired at different times, the batches' samples are not identically distributed. To address this is- sue, we use Group Normalization with 4 groups instead of Batch Normalization in the encoder.

5 | Proposed Framework

To achieve the project goal, the following requirements must be met:

Data acquisition and preprocessing

I. Satellite imagery: Acquire high-resolution satellite images from sources like Sentinel-1, Sentinel-2, Landsat, or MODIS [25].

II. Weather data: Collect meteorological data, including temperature, precipitation, humidity, and wind speed, from weather stations or APIs [26].

III. Soil data: Obtain soil maps and sampling data to characterize soil properties [27].

IV. Data cleaning: Handle missing values, outliers, and inconsistencies in the data.

V. Data preprocessing: Apply image processing techniques (e.g., atmospheric correction, geometric correction) and feature extraction (e.g., vegetation indices, texture analysis) to prepare the data for modeling.

Model development and training

I. Feature engineering: Create relevant features from the preprocessed data, such as vegetation indices, texture features, and temporal trends.

II. Model selection: Choose appropriate machine learning or deep learning algorithms (e.g., random forest, support vector regression, gradient boosting, convolutional neural networks, recurrent neural networks) based on data characteristics and problem complexity.

III. Model training: Train the selected model(s) on the prepared dataset, optimizing hyperparameters for optimal performance.

IV. Model evaluation: Evaluate the trained model(s) using appropriate metrics (e.g., mean squared error, root mean squared error, R-squared) and cross-validation techniques.

Deployment and visualization

I. Deployment: Deploy the trained model(s) to a suitable platform (e.g., cloud-based or local server) for real-time or near-real-time yield predictions.

II. Visualization: Develop user-friendly interfaces (e.g., web applications, dashboards) to visualize the input data, model predictions, and uncertainty estimates.

6 | System Design

Design constraints

I. Data availability: Ensure access to reliable and timely satellite imagery, weather data, and soil information.

II. Computational resources: Utilize sufficient computational resources (e.g., cloud-based computing) to handle large datasets and complex models.

III. Algorithm complexity: Select algorithms that balance accuracy and computational efficiency.

IV. Data quality: Address data quality issues, such as noise, missing values, and outliers.

7 | Experimental Setup

The proposed system architecture consists of the following modules:

Data acquisition and preprocessing module

I. Collect satellite imagery, weather data, and soil data from various sources.

II. Clean and preprocess the data to remove noise, inconsistencies, and missing values.

III. Extract relevant features from the data, such as vegetation indices, texture features, and temporal trends.

Model development and training module

I. Select and train appropriate machine or deep learning models on the preprocessed data.

II. Evaluate the performance of the trained models using appropriate metrics.

III. Optimize hyperparameters to improve model accuracy and generalization.

Deployment and visualization module

I. Deploy the trained model to a suitable platform for real-time or near-real-time predictions.

II. Develop user-friendly interfaces to visualize input data, model predictions, and uncertainty estimates.

Following this system design, we aim to develop a robust and efficient agricultural yield prediction system that can provide valuable insights to farmers, policymakers, and researchers [28].

8 | Experimental Results and Discussion

The panoptic agricultural satellite time series dataset

We introduce PASTIS, the first large-scale, publicly accessible SITS dataset featuring semantic and panoptic annotations[1].

The PASTIS dataset includes 2,433 sequences of multispectral images, each sized 10 x 128 x 128 pixels. Each sequence contains 38 and 61 observations taken from September 2018 to November 2019, totaling over 2 billion pixels.

Acquisition intervals are irregular, averaging 5 days due to automated cloud cover filtering by the satellite provider THEIA. The 10 channels correspond to the Sentinel-2 satellite's non-atmospheric spectral bands, processed with atmospheric correction and resampled to a 10-meter spatial resolution per pixel. This dataset covers roughly 4,000 square kilometers across four distinct regions of France with different climates and crop types, representing about 1% of the French Metropolitan area. Approximately 28% of the images have some degree of cloud cover.

9 | Conclusion

We would like to express our deepest gratitude to all those who supported and guided me throughout the research and paper preparation process.

Firstly, I extend my heartfelt thanks to my supervisor, Dr. Shubadip Parmanik, for their invaluable guidance, constructive feedback, and constant encouragement, which were instrumental in shaping the direction and quality of this study. Their expertise and mentorship have been a cornerstone of our research journey.

We are also grateful to KIIT and the department of computer science for providing the necessary resources and a conducive environment for this work.

We would also like to acknowledge the contributions of our peers and colleagues, whose insightful discussions and camaraderie have greatly enriched this experience.

Thank you all for being part of this journey.

Author Contribution

Nomaan Siraj handled conceptualization, methodology, and how the dataset will be trained and modeled for our research.

Swayam Swaroop validated the results, conducted a formal analysis, and investigated. Anubhav Mishra maintained data, created the initial design, and reviewed and edited the writing.

Tanya Raj did visualization, monitoring, and project management.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Cracknell, A. P. (2007). *Introduction to remote sensing*. CRC press. https://B2n.ir/f88362

[1] This dataset, along with details on its structure, is available at https://github.com/VSainteuf/pastis-benchmark.

[2] Potin, P., Rosich, B., Miranda, N., Grimont, P., Shurmer, I., O'Connell, A., & Gratadour, J. B. (2019). Copernicus sentinel-1 constellation mission operations status. *IGARSS 2019-2019 IEEE international geoscience and remote sensing symposium* (pp. 5385–5388). IEEE. https://doi.org/10.1109/IGARSS.2019.8898949

[3] McCabe, M. F., Houborg, R., & Lucieer, A. (2016). High-resolution sensing for precision agriculture: From earth-observing satellites to unmanned aerial vehicles. *Remote sensing for agriculture, ecosystems, and hydrology XVIII* (pp. 346–355). SPIE. https://doi.org/10.1117/12.2241289

[4] Shoshany, M., Goldshleger, N., & Chudnovsky, A. (2013). Monitoring of agricultural soil degradation by remote-sensing methods: A review. *International journal of remote sensing*, *34*(17), 6152–6181. https://doi.org/10.1080/01431161.2013.793872

[5] Scown, M. W., Brady, M. V, & Nicholas, K. A. (2020). Billions in misspent EU agricultural subsidies could support the sustainable development goals. *One earth*, *3*(2), 237–250. http://dx.doi.org/10.1016/j.oneear.2020.07.011

[6] Halbrendt, J. M., & LaMondia, J. A. (2004). Crop rotation and other cultural practices. In *Nematology: advances and perspectives* (pp. 909-930). Centre for agriculture and biosciences international. https://doi.org/10.1079/9780851996462.0909

[7] Yuan, Y., & Lin, L. (2020). Self-supervised pretraining of transformers for satellite image time series classification. *IEEE journal of selected topics in applied earth observations and remote sensing*, *14*, 474–487. https://doi.org/10.1109/JSTARS.2020.3036602

[8] Guyet, T., & Nicolas, H. (2016). Long term analysis of time series of satellite images. *Pattern recognition letters*, *70*, 17–23. https://doi.org/10.1016/j.patrec.2015.11.005

[9] Kirillov, A., He, K., Girshick, R., Rother, C., & Dollár, P. (2019). Panoptic segmentation. *2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)* (pp. 9404–9413). IEEE. https://doi.org/10.1109/CVPR.2019.00963

[10] Kim, H. S., Song, H., & Jung, J. (2024). Cadastral-to-agricultural: A study on the feasibility of using cadastral parcels for agricultural land parcel delineation. *Remote sensing*, *16*(19), 3568. https://doi.org/10.3390/rs16193568

[11] Kirda, C. (2002). Deficit irrigation scheduling based on plant growth stages showing water stress tolerance. *Food and agricultural organization of the united nations, deficit irrigation practices, water reports*, *22*(102), 3–10. https://www.researchgate.net/publication/284039562

[12] Kim, D., Woo, S., Lee, J. Y., & Kweon, I. S. (2020). *Video panoptic segmentation.* https://doi.org/10.48550/arXiv.2006.11339

[13] Lucas, B., Pelletier, C., Schmidt, D., Webb, G. I., & Petitjean, F. (2020). Unsupervised domain adaptation techniques for classification of satellite image time series. *IGARSS 2020-2020 IEEE international geoscience and remote sensing symposium* (pp. 1074–1077). IEEE. https://doi.org/10.1109/IGARSS39084.2020.9324339

[14] Ye, W., Zhang, W., Lei, W., Zhang, W., Chen, X., & Wang, Y. (2023). Remote sensing image instance segmentation network with transformer and multi-scale feature representation. *Expert systems with applications*, *234*, 121007. https://doi.org/10.1016/j.eswa.2023.121007

[15] Heidarianbaei, M., Kanyamahanga, H., & Dorozynski, M. (2024). Temporal ViT-U-Net tandem model: Enhancing multi-sensor land cover classification through transformer-based utilization of satellite image time series. *International society for photogrammetry and remote sensing annals of the photogrammetry, remote sensing and spatial information sciences*, *10*, 169–177. https://doi.org/10.5194/isprs-annals-X-3-2024-169-2024

[16] Chuang, Y., Zhang, S., & Zhao, X. (2023). Deep learning-based panoptic segmentation: Recent advances and perspectives. *Institution of engineering and technology image processing*, *17*(10), 2807–2828. https://doi.org/10.1049/ipr2.12853

[17] Elharrouss, O., Al-Maadeed, S., Subramanian, N., Ottakath, N., Almaadeed, N., & Himeur, Y. (2021). *Panoptic segmentation: A review*. https://doi.org/10.48550/arXiv.2111.10250

[18] Lee, Y., & Park, J. (2020). Centermask: Real-time anchor-free instance segmentation. *2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR)* (pp. 13906–13915). IEEE. https://doi.org/10.1109/CVPR42600.2020.01392

[19] Singh, Y., Bhatia, P. K., & Sangwan, O. (2007). A review of studies on machine learning techniques. *International journal of computer science and security*, *1*(1), 70–84. https://www.researchgate.net/profile/Pradeep-Bhatia-2/publication/41845861

[20] Poria, S., Cambria, E., & Gelbukh, A. (2016). Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-based systems*, *108*, 42–49. https://doi.org/10.1016/j.knosys.2016.06.009

[21] Chennupati, S., Narayanan, V., Sistu, G., Yogamani, S., & Rawashdeh, S. A. (2021). Learning panoptic segmentation from instance contours. *2021 IEEE international conference on robotics and automation (ICRA)* (pp. 9586–9593). IEEE. https://doi.org/10.1109/ICRA48506.2021.9560798

[22] Petitjean, C., & Dacher, J. N. (2011). A review of segmentation methods in short axis cardiac MR images. *Medical image analysis*, *15*(2), 169–184. https://doi.org/10.1016/j.media.2010.12.004

[23] Roscher, R., Volpi, M., Mallet, C., Drees, L., & Wegner, J. D. (2020). Semcity toulouse: A benchmark for building instance segmentation in satellite images. *International society for photogrammetry and remote sensing annals of the photogrammetry, remote sensing and spatial information sciences*, *5*, 109–116. https://doi.org/10.5194/isprs-annals-V-5-2020-109-2020

[24] Wang, Y. L., Li, L., Zhou, S. W., & Chen, X. (2024). A multi-center polarmask model for image segmentation. *Fourth symposium on pattern recognition and applications (SPRA 2023)* (pp. 62–70). SPIE. https://doi.org/10.1117/12.3030078

[25] Burke, M., Driscoll, A., Lobell, D. B., & Ermon, S. (2021). Using satellite imagery to understand and promote sustainable development. *Science*, *371*(6535), eabe8628. https://doi.org/10.1126/science.abe8628

[26] Herrera, M., Natarajan, S., Coley, D. A., Kershaw, T., Ramallo-González, A. P., Eames, M., … ., & Wood, M. (2017). A review of current and future weather data for building simulation. *Building services engineering research and technology*, *38*(5), 602–627. https://doi.org/10.1177/0143624417705937

[27] Shangguan, W., Dai, Y., Duan, Q., Liu, B., & Yuan, H. (2014). A global soil data set for earth system modeling. *Journal of advances in modeling earth systems*, *6*(1), 249–263. https://doi.org/10.1002/2013MS000293

[28] Gupta, A., & Nahar, P. (2023). Classification and yield prediction in smart agriculture system using IoT. *Journal of ambient intelligence and humanized computing*, *14*(8), 10235–10244. https://doi.org/10.1007/s12652-021-03685-w%0A%0A