# UNet Segmentation for Accurate Parcel Delineation: Improvement with temporal multi-spectral images

Amine Hadir [1] , Mohamed Adjou [1], Gaëtan Palka[2], Olga Assainova[3], Marwa El Bouz[3]

[1] Laboratoire ISEN, *ISEN Brest*
[2] Centre d'Etudes pour le Développement des Territoires et l'Environnement, *Université d'Orléans*
[3] Laboratoire ISEN, *ISEN Brest*

**Keywords:** Parcel Delineation, Temporal Multi-spectral dataset, Remote Sensing, Deep learning.

**Abstract**

Agricultural parcel delineation is critical for generating cadastral maps that underpin sustainable land management, precision agriculture, and data-driven policymaking. While satellite imaging provides a scalable solution, most existing approaches rely on static RGB or single-date spectral data, neglecting the temporal dynamics of agricultural landscapes. This study introduces **TempAgriBound**, a novel temporal multispectral dataset designed to advance parcel boundary detection by capturing both spectral and phenological features. The dataset comprises a dense time-series of Sentinel-2 multispectral imagery (10 bands at 10m resolution) and derived spectral indices (NDVI, NDWI, SAVI, etc.), spanning the entire 2023 growing season in Brittany, France (a region characterized by diverse crop rotations and fragmented landholdings). We propose a 3D U-Net architecture optimized for spatiotemporal feature extraction, which processes multi-spectral time stacks to exploit crop growth stages and seasonal spectral variations. For comparison, a 2D U-Net baseline using single-date RGB composites was implemented. By systematically evaluating these models, we aim to determine the differential effects of temporal spectral information on parcel boundary detection. These findings underscore the synergistic value of temporal resolution and spectral diversity in automated parcel mapping, particularly in regions with complex crop patterns. The study advances scalable precision agriculture tools and provides actionable insights for integrating temporal-spectral data into national land registries.

## 1. Introduction

Cartographic delimitation of agricultural land parcels historically represented a complex and resource-intensive cartographical challenge based on manual geospatial measurement methodologies. Before the emergence of computational visual recognition and machine learning algorithmic approaches, agricultural land boundary delineation required extensive human intervention and sophisticated geodetic instrumentation, including precision technologies such as theodolites, total stations, and Global Positioning System (GPS) receivers (van der Molen, 2002).

Remote Sensing Images (RSI), with its improved spatial and spectral resolution capabilities alongside consistent temporal sampling frequencies, is vital in providing the required multi-dimensional information for this objective. Modern satellite constellations deliver systematic revisit capabilities ranging from daily to biweekly coverage, creating dense time series that capture both abrupt and gradual boundary transitions. Spectral information, captured across multiple wavelength bands, offers unique insights by differentiating vegetation types, crop health, soil moisture, and land use characteristics, thereby enabling more precise and automated boundary identification beyond traditional visual mapping techniques (Yu et al., 2024). When combined with multi-temporal analysis frameworks, these spectral signatures can be traced through time to reveal persistent boundary features while filtering ephemeral patterns, significantly enhancing delineation accuracy and consistency.

The diversity in agricultural parcel complexity presents a unique challenge in the field. In Europe, particularly in countries like France, agricultural landscapes are characterized by intricate patterns of fragmented parcels, shaped by centuries of inheritance laws and traditional farming practices (Bartels et al., 2024). This contrasts sharply with regions such as North America or Australia, where industrial farming techniques have led to larger and more consolidated parcels. These regional variations, rooted in historical, legal and geographical factors, significantly influence the effectiveness of parcel delineation methods and underscore the importance of developing robust and adaptable approaches.

The temporal dimension adds critical depth to this analysis, such as boundary delineation, crop rotation patterns, and long-term land use transitions occurring at multiple timescales. Agricultural boundaries experience both cyclical changes (seasonal crop patterns) and linear transformations (gradual field expansions or consolidations), creating a four-dimensional problem that static mapping approaches fail to address (Aung et al., 2020). When properly quantified, this temporality can be leveraged as a force multiplier in boundary detection accuracy, enabling detection of subtle boundary shifts that would otherwise remain imperceptible in single-timepoint analyses.

In the evolving technical landscape, deep convolutional neural networks (CNNs) have revolutionized the field of computer vision, demonstrating unprecedented performance across diverse applications including image classification (Adegun et al., 2023, Gupta et al., 2021), object detection (Gui et al., 2024, Zhang et al., 2023), semantic segmentation (Hadir et al., 2024, Li et al., 2024), and land cover/land use mapping (Sertel et al., 2022, Fayaz et al., 2024). The domain of agricultural parcel delineation has been particularly transformed by these advances, where deep learning methodologies have exhibited exceptional prowess in automatically extracting and defining field boundaries through sophisticated learning of multidimensional representations, spectral signatures, textural patterns, temporal dynamics, and spatial relationships, from remote sensing imaging (RSI).

The architectural evolution in this domain began with the groundbreaking U-Net (Ronneberger et al., 2015) framework, which

established fundamental encoding-decoding principles with skip connections. This foundation was subsequently refined through architectures like ResUNet-a (Wagner et al., 2020), which incorporated critical innovations including atrous convolutions for multi-scale feature extraction and pyramid scene parsing pooling to enhance contextual awareness. Authors propose new derivatives of ResUNet architectures are Recurrent Residual UNet, Adversarial ResUNet. For the Recurrent Residual UNet, we find recurrent connections in a neural network allow the network to maintain a form of "memory" by using its output as an additional input for the next step. This is particularly useful when dealing with sequential data, where the order of data points is important. In the context of image segmentation, recurrent connections can help the model capture temporal dependencies in image sequences, such as changes in an agricultural field over time. The progression continued with the introduction of computationally efficient FracTALResUNet (Waldner et al., 2021), which implemented blocks inspired by the fractal network (**?**) as a backbone, that dramatically improved feature extraction capabilities while maintaining computational efficiency. Complementing these advances, object-centric frameworks such as the adapted Mask R-CNN (He et al., 2017, Meyer et al., 2020) have pushed boundaries by elegantly integrating instance detection with pixel-precise segmentation, resulting in closed topologically sound geometries that accurately represent field boundaries in real-world agricultural landscapes.

Temporal dynamics has been addressed through cutting-edge segmentation approaches that take advantage of satellite image time series (SITS) to captute the subtle but distinctive crop phenology patterns that emerge over growing seasons (Garnot and Landrieu, 2021). These temporal models excel at distinguishing adjacent fields with similar crops at single time points but divergent developmental trajectories.

Recently, the latest research is based on Transformers (Vaswani et al., 2023), which have been shown to perform well with satellite images for various tasks. This study (Aleissaee et al., 2023) showed that more peer-reviewed studies used transformers in Remote Sensing (RS). This paper (Xu et al., 2023) proposes a new multi-swin mask transformer (MSMTransformer) method based on Vision Transformers (ViT) as backbone phase. In addition, it's worth mentioning the famous Transformers-based model, SAM (Segment Any Things) (Kirillov et al., 2023). In addition, authors in (Hadir et al., 2025) present a fine-tuning of LoRA-SAM for agricultural parcel delineation.

The paradigm shift in agricultural parcel mapping methodologies has catalyzed a fundamental transformation in how we conceptualize, implement, and automate boundary delineation processes across diverse agricultural landscapes. These advancements address the multifaceted challenges inherent in agricultural land monitoring while establishing new benchmarks for accuracy, efficiency, and scalability in boundary detection systems. Our research focuses specifically on evaluating the critical role of temporal spectral information dimensions in agricultural field boundary segmentation tasks. To accomplish this, we designed a systematic comparative framework utilizing variations of the established U-Net architecture—specifically contrasting a standard 2D U-Net implementation optimized for RGB imagery against a temporally-aware 3D U-Net variant engineered to process multi-spectral data streams.

Our selection of U-Net as the foundational architecture for this investigation is strategically justified by dual considerations: first, its well-established position within the field, evidenced by numerous successful implementations and architectural derivatives specifically tailored for agricultural segmentation challenges (Ronneberger et al., 2015, Wagner et al., 2020, Waldner et al., 2021); and second, its architectural transparency, which provides an ideal experimental tested for isolating and quantifying the specific contributions of temporal spectral information processing without the confounding influence of overly complex architectural components. Furthermore, this choice is further supported by the recent study (Hadir et al., 2025), in which deep learning models for agricultural parcel delineation are categorized into different categories, with particular emphasis placed on U-Net and its derivatives due to their strong performance and adaptability across varying levels of parcel complexity.

## 2. Methods

### 2.1 Data description and generation

**TempAgriBound** as shown in Figure1, is a multi-temporal and multi-spectral dataset constructed from Sentinel-2 RSI collected on three distinct dates in 2023: February 15th, June 15th, and October 8th. Spectral bands B05, B06, B07, B011 and B012 with 20-meter spatial resolution were resample at 10 meter using B-Spline interpolation.

The parcel boundaries were derived from the official french Graphical Parcel Register (GPR) of 2023. These boundaries, representing agricultural plot limits, rasterised to match the spatial parameters of the RSI — including the 10-meter spatial resolution, coordinate reference system (CRS), and geographic extent.

The parcel boundaries were also converted into vector polygons to create parcels and then rasterised as above to be combined with a cloud mask was generated using the Scene Classification Layer (SCL) provided by Sentinel-2. As the SCL is available at 20m resolution, it was resampled to 10m using nearest-neighbor interpolation to align with the resolution of the multispectral bands. Only cloud-free areas were retained in the dataset.

Specifically, image patches where at least 87classified as cloud-free parcels and where the top-left corner reference pixels were spaced at least 20 pixels apart were selected. This spatial separation criterion was applied to ensure diverse, non-overlapping samples throughout the dataset.

The final dataset consists of 4,005 pairs of samples, each consisting of a label file (limits_X_Y) and a multispectral remote sensing image (remote_X_Y). Each RSI patch (remote_X_Y) is a 3-date temporal stack with the shape (3, 9, 224, 224), where:

- The first dimension represents the three acquisition dates (February, June, and October),
- The second dimension corresponds to nine Sentinel-2 bands: B02 (490nm, Blue), B03 (560nm, Green), B04 (665nm, Red), B05, B06 and B07 (705nm, 740nm and 783nm, Red Edge bands), B08 (842nm, NIR), B11 and B12 (1,610nm and 2,190nm, SWIR)
- The final two dimensions are the spatial size of each thumbnail, covering 2.24 km × 2.24 km at a 10-meter resolution.

This dataset provides a rich temporal and spectral representation of vegetation dynamics across agricultural parcels and is particularly suited for tasks such as crop classification, vegetation health assessment, and change detection. Indeed, it gives access to a wide range of wavelengths, especially near infrared and short-wave infrared, that allow the construction of indicators as NDVI.

### 2.2 Neural Network architecture

UNet architectures (U-Net2D for RGB and vegetation bands (B04, B08 and B11), while U-Net3D for temporal multispectral imagery) were selected as baseline models for agricultural boundary segmentation due to their architectural transparency and proven effectiveness. This deliberate selection of streamlined network designs facilitates our primary research goal: to isolate and rigorously evaluate the intrinsic capabilities of spectro-temporal information in agricultural parcel delineation, without architectural complexities masking these fundamental contributions. The temporal-spectral dimension introduces critical information by capturing phenological transitions across growing seasons, revealing boundary distinctions through differential vegetation response patterns that remain imperceptible in single-timepoint analyses.

The U-Net's elegant encoder-decoder structure with skip connections provides an ideal experimental platform to examine how the fusion of temporal dynamics with spectral signatures can significantly enhance boundary detection performance. By processing both the spatial relationships between pixels and the characteristic spectral reflectance patterns across multiple wavelengths and time points, we can identify distinctive temporal-spectral signatures that emerge at field boundaries. These signatures often manifest as discontinuities in crop growth cycles, management practices, and soil characteristics that become pronounced when analyzed across the temporal dimension. This methodological approach ensures that performance improvements can be directly attributed to the spectro-temporal information's inherent discriminative power, unobscured by complex neural network architectures.

All UNet implementations were developed using PyTorch framework. Training was conducted with Adam Optimizer at a learning rate of 1e-3, utilizing a batch size of 8 across 100 epochs on NVIDIA TITAN V GPU hardware. Binary cross entropy (BCE) served as the loss function for all model variants. The dataset underwent random distribution splitting with proportions of 80/10/10 for training, validation, and testing respectively.

### 2.3 Metrics

In the context of land parcel delineation, evaluating methods performance is important and a crucial point. Whereas global metrics such as accuracy, F1 score and mIoU assess global performance, boundary metrics like BDE (Boundary Displacement Error) focusing on pixels close to the limits can offer more relevant information about method performances in the case of parcel delineation(Unnikrishnan et al., 2007).

BDE measures specifically the average displacement between the predicted parcel boundary and the ground truth boundary. It quantifies the spatial accuracy of the delineated parcel boundaries. Lower BDE values indicate a closer match between predicted and actual boundaries, which means better precision in boundary delineation.

To accomplish this, evaluation metrics are utilized to measure accuracy and effectiveness. We evaluated all models in terms of accuracy, F1 score, mIoU (mean intersection over union) and BDE (boundary displacement error) metrics.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \quad (2)$$

$$JaccardIndex = \frac{TP}{TP + FP + FN} \quad (3)$$

Or,

$$IoU = \frac{Intersection}{Union} \quad (4)$$

Where, True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

$$BDE = \frac{1}{2}\left(\frac{1}{|B_s|}\sum_{ps_s \in B_s}\sum_{p_gt \in B_gt} min\{d_E(ps_s, p_gt)\} + \frac{1}{|B_gt|}\sum_{p_gt \in B_gt}\sum_{ps_s \in B_s} min\{d_E(p_gt, ps_s)\}\right), \quad (5)$$

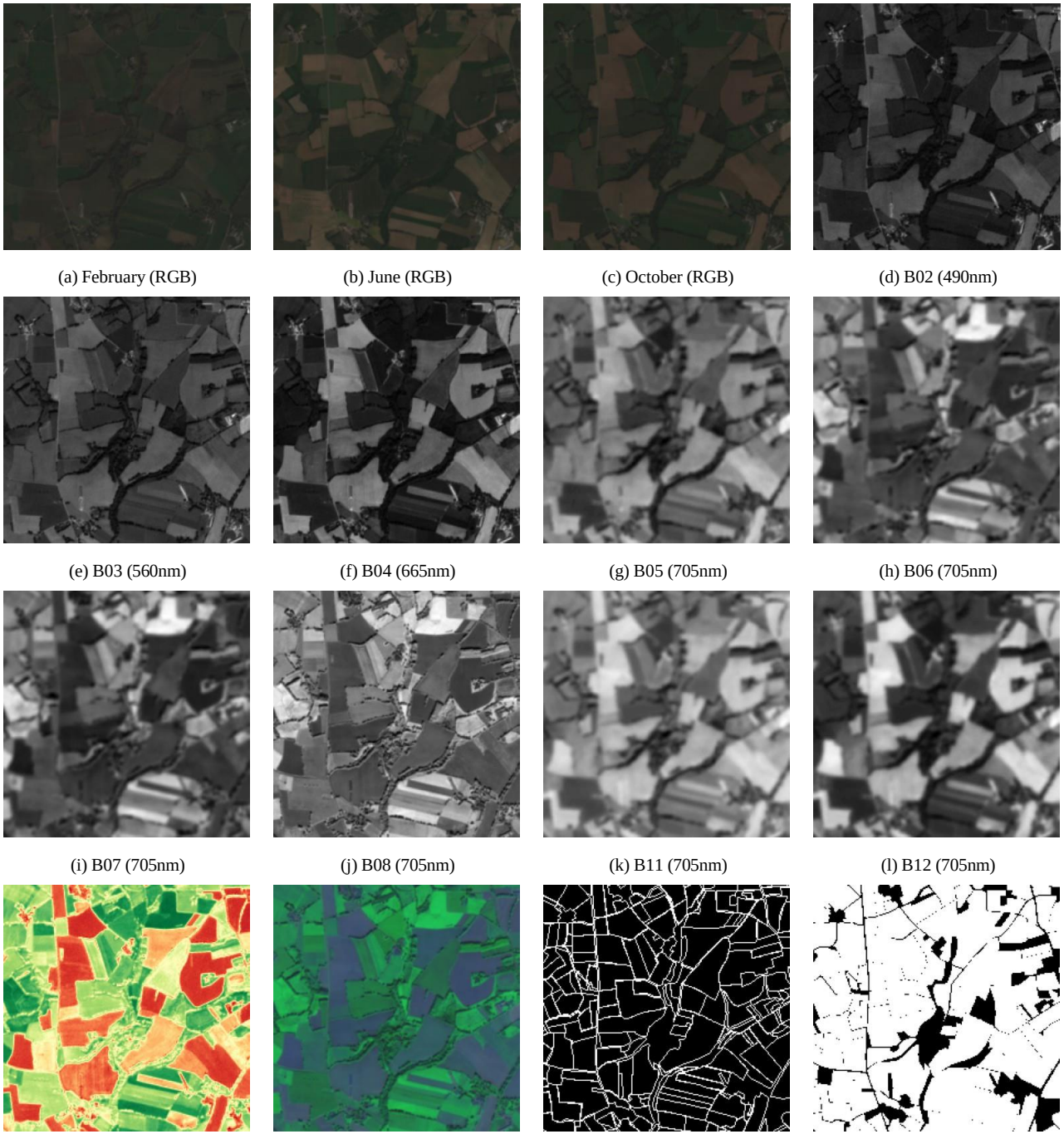
Where:

- the distance ($d_E$) between a boundary pixel ($b_s$) in the obtained boundary image ($B_s$) and the closest pixel ($b_gt$) in the ground truth boundary image ($B_gt$) is used to define the disagreement (error) of each boundary pixel.
- The lower the BDE value, the more accurate the segmentation results. By minimizing BDE, parcel delineation algorithms can improve the accuracy and reliability of the resulting parcel boundaries (Garcia-Pedrero et al., 2019).

## 3. Results & Discussion

TThe experimental results presented in Table 1 demonstrate the significant impact of both architectural choices and input data configurations on agricultural parcel delineation performance. Our analysis reveals clear patterns across the different U-Net variants tested. The 3D U-Net models consistently outperform their 2D counterparts across all evaluation metrics by a substantial margin. The four 3D implementations achieve F1-Scores ranging from 0.80-0.85 compared to 0.42-0.63 for 2D models—a performance improvement of approximately 35-50%. This substantial difference highlights the critical importance of the 3D architecture's ability to effectively process spatial information.

Among the 3D implementations, the variant using temporal and spectral dimensions achieves exceptional boundary delineation with a minimal BDE of 0.17 alongside strong segmentation metrics (F1-Score: 0.84, Accuracy: 0.95). Similarly, the U-Net3D processing temporal and MSI data shows excellent performance with the highest F1-Score (0.85) while maintaining



(a) February (RGB) (b) June (RGB) (c) October (RGB) (d) B02 (490nm)

(e) B03 (560nm) (f) B04 (665nm) (g) B05 (705nm) (h) B06 (705nm)

(i) B07 (705nm) (j) B08 (705nm) (k) B11 (705nm) (l) B12 (705nm)

(m) NDVI (n) Veg bands (B04,B08,B11) (o) Parcels limits (p) Cloudless parcels

Figure 1. A sample of patch with the RGB image for the three dates, all bands, NDVI and veg. bands for october, and its corresponding parcel limits used in the experiences

a competitive BDE (0.18). Interestingly, the U-Net3D using only MSI data without temporal information still performs remarkably well (F1-Score: 0.84, Accuracy: 0.95), challenging our initial assumption about the critical importance of temporal data. This suggests that the architectural advantages of 3D U-Net for spatial processing may be more significant than previously thought, even when temporal information is absent.

For 2D architectures, we observe that input band selection significantly impacts performance. The configuration incorporating vegetation bands with NDVI demonstrates the strongest performance among 2D variants (F1-Score: 0.63, mIoU: 0.70), suggesting that NDVI provides valuable discriminative information. In contrast, the RGB-based U-Net2D model shows poor boundary precision with the highest BDE (1.57) despite reasonable overall accuracy (0.80), indicating a tradeoff between global classification and boundary precision. The U-Net2D with spectral dimension performs particularly poorly in accuracy (0.063), suggesting potential issues with this specific configuration.

The technical necessity of reshaping multi-dimensional inputs for U-Net2D models (from (3,4,224,224) to (3*4,224,224) for vegetation bands + NDVI) likely contributes to their reduced effectiveness, as this flattening may disrupt the natural relation-

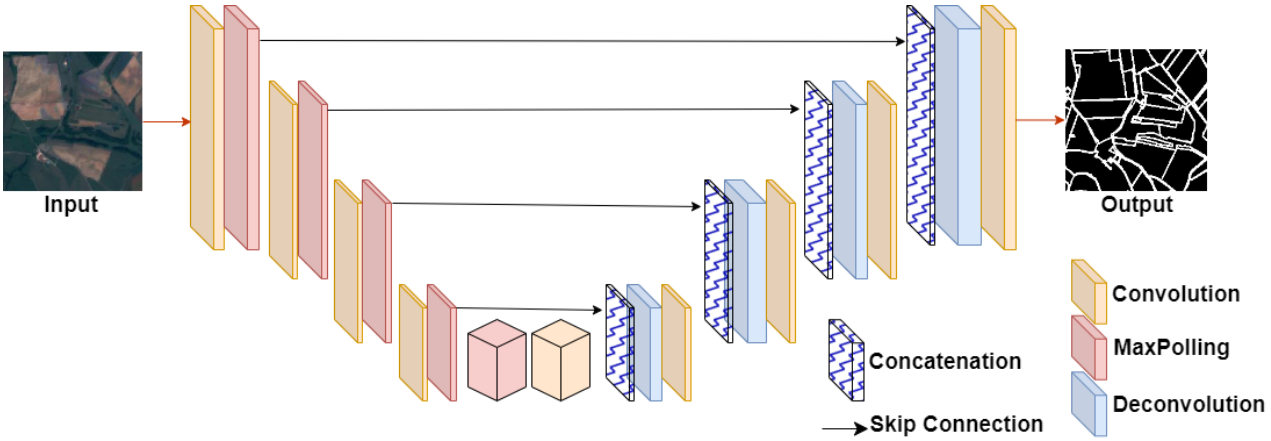



Figure 2. U-Net2D architectures

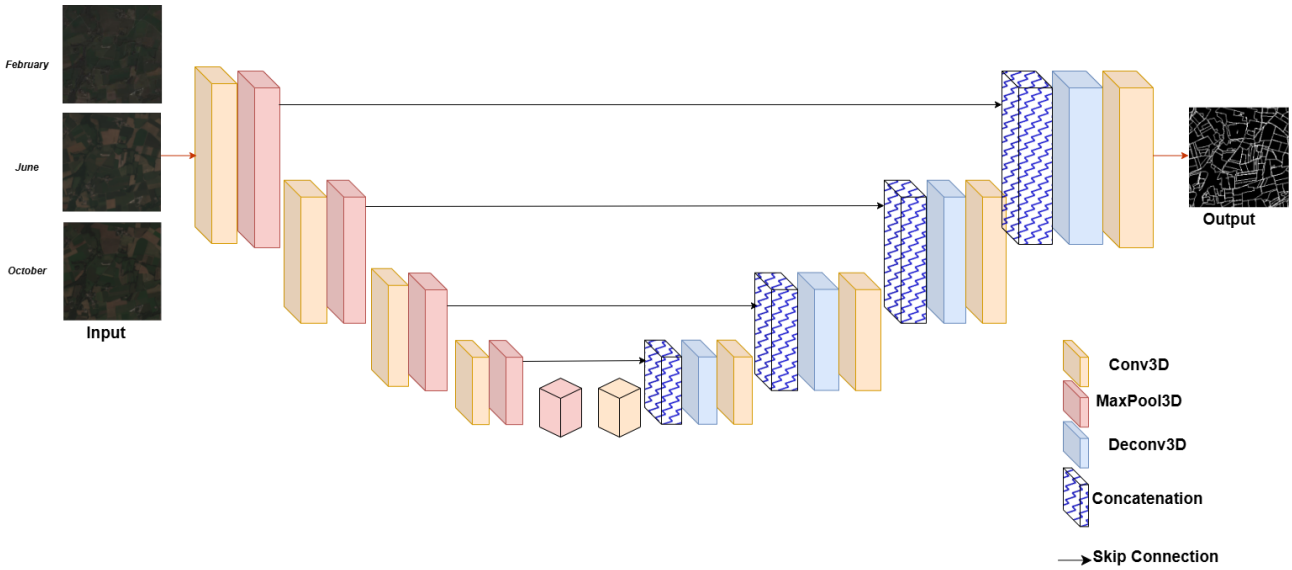



Figure 3. U-Net3D architectures

| Method | BDE | mIoU | F1-Score | Accuracy |
|---|---|---|---|---|
| U-Net2D [stack temporal and RGB dim] | 1.57 | 0.40 | 0.50 | 0.80 |
| U-Net2D [stack temporal and spectral dim] | 0.89 | 0.62 | 0.46 | 0.0.63 |
| U-Net2D [stack temporal and veg bands: B04, B08, B11] | 0.89 | 0.59 | 0.42 | 0.69 |
| U-Net2D [stack temporal and veg bands (add NDVI)] | 0.88 | 0.70 | 0.63 | 0.70 |
| U-Net3D temporal and veg bands (add NDVI) | 0.26 | 0.67 | 0.80 | 0.93 |
| U-Net3D temporal and MSI | 0.18 | 0.70 | 0.85 | 0.97 |
| U-Net3D [stack temporal and spectral dim] | 0.17 | 0.73 | 0.84 | 0.95 |
| U-Net3D MSI no tempral | 0.25 | 0.73 | 0.84 | 0.95 |

Table 1. Performance comparison of U-Net variants for parcel delineation, evaluated using metrics such as Accuracy, F1-Score, Mean Intersection over Union (mIoU), and Boundary Displacement Error (BDE) across different dataset types.

ship between spectral bands. Conversely, 3D U-Net architectures can directly process the data in its native multi-dimensional structure, preserving these important spatial relationships. Boundary Displacement Error values demonstrate that 3D architectures (BDE range: 0.17-0.26) achieve substantially more precise boundary localization compared to 2D approaches (BDE range: 0.88-1.57), regardless of whether temporal information is included.

The superior performance of 3D U-Net variants can be attributed to their ability to directly model spatial relationships rather than flattening dimensions. While we previously emphasized the importance of temporal modeling, the strong performance of the non-temporal 3D MSI model suggests that the architectural benefits of 3D convolutions extend beyond temporal modeling to better spatial feature extraction. The U-Net architecture's skip connections effectively transfer spatial details in both implementations, but the 3D variants appear to maintain better context through these connections.

Our findings demonstrate that the architectural capacity to process spatial relationships in their native three-dimensional structure represents the dominant factor determining overall performance in agricultural parcel delineation. While input band selection produces measurable variations within each architectural group (with vegetation bands + NDVI consistently outperforming other band combinations in 2D models), the dimensional handling of data creates a more substantial performance gap between model families. The comparable performance between temporal and non-temporal 3D models suggests that future research should carefully evaluate whether the increased complexity of temporal data processing is justified by performance gains in specific agricultural monitoring applications.

## 4. Conclusion

This study presents **TempAgriBound**, a novel temporal multispectral dataset for agricultural parcel delineation, along with comprehensive performance evaluations of 2D and 3D U-Net architectures. Our experiments demonstrate that incorporating temporal information significantly enhances boundary detection accuracy in agricultural settings, with 3D U-Net models consistently outperforming their 2D counterparts across all evaluation metrics.

The experimental results clearly illustrate that the 3D U-Net architecture leveraging both temporal and spectral dimensions achieves superior performance with a minimal Boundary Displacement Error of 0.17 and strong segmentation metrics (F1-Score of 0.84, Accuracy of 0.95). This marked improvement over the 2D U-Net variants (where the best model achieved only 0.63 F1-Score with a BDE of 0.88) underscores the critical importance of capturing phenological changes throughout the growing season for precise parcel delineation.

We observe that while spectral band selection contributes to performance variations, the architectural capacity to process temporal relationships represents the dominant factor determining overall performance. The substantial performance gap between 2D and 3D architectures (with 3D models achieving F1-Scores approximately 20-40% higher) confirms our hypothesis that temporal dynamics provide essential contextual information that significantly enhances both overall segmentation quality and boundary precision.

The TempAgriBound dataset and our optimized 3D U-Net approach offer several practical advantages for agricultural monitoring systems. First, they enable more accurate cadastral mapping in regions with complex crop rotations and fragmented landholdings, such as Brittany, France. Second, the improved boundary detection precision supports fine-grained analysis for precision agriculture applications. Finally, the methodology's ability to leverage freely available Sentinel-2 imagery ensures scalability and accessibility across diverse agricultural landscapes globally. Future work should explore the transferability of our approach to other regions with different agricultural practices and seasonal patterns.

Additionally, incorporating auxiliary data sources such as digital elevation models or soil maps could further enhance boundary detection in challenging scenarios where spectral-temporal information alone may be insufficient.

In conclusion, this study advances the field of agricultural parcel delineation by demonstrating the synergistic value of temporal resolution and spectral diversity in automated mapping. The proposed methodology provides a foundation for developing more robust land monitoring systems that can support sustainable agricultural practices, environmental conservation efforts, and evidence-based policy decisions.

## References

Adegun, A. A., Viriri, S., Tapamo, J.-R., 2023. Review of deep learning methods for remote sensing satellite images classification: experimental survey and comparative analysis. 10(1), 93. https://doi.org/10.1186/s40537-023-00772-x.

Aleissaee, A. A., Kumar, A., Anwer, R. M., Khan, S., Cholakkal, H., Xia, G.-S., Khan, F. S., 2023. Transformers in Remote Sensing: A Survey. *Remote Sensing*, 15(7). https://www.mdpi.com/2072-4292/15/7/1860.

Aung, H. L., Uzkent, B., Burke, M., Lobell, D., Ermon, S., 2020. Farm Parcel Delineation Using Spatio-temporal Convolutional Networks. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Seattle, WA, USA, 340–349.

Bartels, C., Ja¨ger, S., Obergruber, N., 2024. Long-Term Effects of Equal Sharing: Evidence from Inheritance Rules for Land. *The Economic Journal*, ueae040. https://doi.org/10.1093/ej/ueae040. eprint: https://academic.oup.com/ej/advance-article-pdf/doi/10.1093/ej/ueae040/59248160/ueae040.pdf.

Fayaz, M., Nam, J., Dang, L. M., Song, H.-K., Moon, H., 2024. Land-Cover Classification Using Deep Learning with High-Resolution Remote-Sensing Imagery. *Applied Sciences*, 14(5). https://www.mdpi.com/2076-3417/14/5/1844.

Garcia-Pedrero, A., Lillo-Saavedra, M., Rodriguez-Esparragon, D., Gonzalo-Martin, C., 2019. Deep Learning for Automatic Outlining Agricultural Parcels: Exploiting the Land Parcel Identification System. *IEEE Access*, 7, 158223–158236.

Garnot, V. S. F., Landrieu, L., 2021. Panoptic Segmentation of Satellite Image Time Series with Convolutional Temporal Attention Networks. *CoRR*, abs/2107.07933. https://arxiv.org/abs/2107.07933.

Gui, S., Song, S., Qin, R., Tang, Y., 2024. Remote Sensing Object Detection in the Deep Learning Era—A Review. *Remote Sensing*, 16(2). https://www.mdpi.com/2072-4292/16/2/327.

Gupta, S., Dwivedi, R. K., Kumar, V., Jain, R., Jain, S., Singh, M., 2021. Remote sensing image classification using deep learning. *2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART)*, 274–279.

Hadir, A., Adjou, M., Assainova, O., Palka, G., Elbouz, M., 2025. Comparative study of agricultural parcel delineation deep learning methods using satellite images: Validation through parcels complexity. *Smart Agricultural Technology*, 10, 100833. https://www.sciencedirect.com/science/article/pii/S2772375525000662.

Hadir, A., Assainova, O., Adjou, M., Palka, G., Elbouz, M., Al Falou, A., 2024. Bridging the gap: deep learning in the semantic segmentation of remote sensing data. *Pattern Recognition and Tracking XXXV*, 13040, SPIE, 182–191.

He, K., Gkioxari, G., Dolla´r, P., Girshick, R., 2017. Mask r-cnn. *Proceedings of the IEEE international conference on computer vision*, 2961–2969.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y. et al., 2023. Segment anything. *arXiv preprint arXiv:2304.02643*.

Li, J., Cai, Y., Li, Q., Kou, M., Zhang, T., 2024. A review of remote sensing image segmentation by deep learning methods. *International Journal of Digital Earth*, 17(1), 2328827. https://doi.org/10.1080/17538947.2024.2328827.

Meyer, L., Lemarchand, F., Sidiropoulos, P., 2020. A DEEP LEARNING ARCHITECTURE FOR BATCH-MODE FULLY AUTOMATED FIELD BOUNDARY DETECTION. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B3-2020, 1009–1016.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597 [cs].

Sertel, E., Ekim, B., Ettehadi Osgouei, P., Kabadayi, M. E., 2022. Land Use and Land Cover Mapping Using Deep Learning Based Segmentation Approaches and VHR Worldview-3 Images. *Remote Sensing*, 14(18). https://www.mdpi.com/2072-4292/14/18/4558.

Unnikrishnan, R., Pantofaru, C., Hebert, M., 2007. Toward Objective Evaluation of Image Segmentation Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6), 929-944.

van der Molen, P., 2002. Property rights, land registration and cadastre in the european union.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I., 2023. Attention is all you need.

Wagner, M., Oppelt, Natascha, 2020. Deep learning and adaptive graph-based growing contours for agricultural field extraction. *Remote sensing*, 12(12), 1990.

Waldner, F., Diakogiannis, F. I., Batchelor, K., Ciccotosto-Camp, M., Cooper-Williams, E., Herrmann, C., Mata, G., Toovey, A., 2021. Detect, Consolidate, Delineate: Scalable Mapping of Field Boundaries Using Satellite Images. *Remote Sensing*, 13(11), 2197.

Xu, Y., Xue, X., Sun, Z., Gu, W., Cui, L., Jin, Y., Lan, Y., 2023. Deriving Agricultural Field Boundaries for Crop Management from Satellite Images Using Semantic Feature Pyramid Network. *Remote Sensing*, 15(11). https://www.mdpi.com/2072-4292/15/11/2937.

Yu, Z., Li, T., Zhu, Y., Pan, R., 2024. Exploring foundation models in remote sensing image change detection: A comprehensive survey.

Zhang, X., Zhang, T., Wang, G., Zhu, P., Tang, X., Jia, X., Jiao, L., 2023. Remote sensing object detection meets deep learning: A metareview of challenges and advances. *IEEE Geoscience and Remote Sensing Magazine*, 11(4), 8–44.