IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 3 July 2025, accepted 24 September 2025, date of publication 20 November 2025, date of current version 5 December 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3626678

RESEARCH ARTICLE

Multitemporal Multispectral Dataset for Palaeochannels Segmentation (MAPS)

GIULIO POGGI[1], ANDALEEB YASEEN[1,2], RAVEERAT JATURAPITPORNCHAI[1], SARA FERRO[1], GREGORY SECH[1], PETER NAYLOR[3], MARIA CRISTINA SALVI[1], SEBASTIANO VASCON[2], BERTRAND LE SAUX[3,4], (Senior Member, IEEE), MARCO FIORUCCI[1], AND ARIANNA TRAVIGLIA[1], (Associate Member, IEEE)

[1]Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, 31056 Treviso, Italy
[2]Department of Environmental Sciences, Informatics and Statistics, Università Ca' Foscari, 30171 Venezia, Italy
[3]Φ-Laboratory, European Space Agency, 00044 Frascati, Italy
[4]AI4EARTH, 31000 Toulouse, France

Corresponding author: Arianna Traviglia (arianna.traviglia@iit.it)

This work was supported by the European Space Agency (ESA) through the "Cultural Landscapes Scanner" Project under Contract 4000132058.

ABSTRACT Deep learning methods have significantly advanced the analysis of satellite imagery, greatly enhancing the efficiency of screening extensive datasets. However, progress in the automatic identification of elements under the surface (subsoil features), detectable only through proxies like soil and vegetation alterations, has been limited. This challenge is further compounded by the significant influence of seasonal variability and land use management on their visibility. Automatic identification often requires multitemporal observations to maximise the likelihood of obtaining favourable visibility conditions and to strengthen the reliability of each identification. Additionally, the limited availability of publicly accessible specialised datasets has impeded progress in applying deep learning approaches. This paper presents the first publicly available dataset of multitemporal Sentinel-2 imagery for identifying palaeochannels, a specific category of subsoil features. We established a baseline for semantic segmentation by evaluating three architectures with various parameter configurations, specifically designed to address the unique challenges of segmenting these elongated and branched features under diverse landscape conditions at different times of the year. The results offer preliminary insights into addressing the varying visibility of subsoil features by applying deep learning to analyse temporal series. The dataset is available on Hugging Face and IIT Dataverse repositories.

INDEX TERMS Benchmark dataset, multispectral images, palaeochannels, semantic segmentation.

I. INTRODUCTION

Optical imagery-based remote sensing has been extensively used to detect subsoil anthropogenic or environmental features by identifying surface 'markers' on soils and vegetation that suggest the presence of underlying materials or conditions. The detection of these proxies has been instrumental in pinpointing subsoil archaeological and geomorphological features, among others [1], [2], [3], [4]. In arable land with low or absent vegetation, the visibility of subsoil features can be hindered by several factors, including soil properties (such as moisture content, texture, and composition), agricultural practices, seasonal climate variations, and obstruction caused by fragmentation of the agricultural landscape and modern infrastructures. Given these variables, best remote sensing-based subsoil feature detection practices typically involve screening the same area at different times throughout the year. This approach utilises repeated observations to enhance the reliability of each identification and maximise the likelihood of capturing favourable visibility conditions [5], [6], [7], [8]. However, its focus on localised, landscape-specific factors limits broader applicability, as variations in climate, soil composition, and land use across different regions affect the transferability of results between different study areas. Despite advancements in Earth Observation cloud computing platforms (e.g. Google Earth

The associate editor coordinating the review of this manuscript and approving it for publication was Geng-Ming Jiang.

Engine, Sentinel Hub), which enable the processing of extensive satellite image catalogues, the selection of individual images under optimal conditions remains a preferred strategy for detecting subtle or short-lived traces [9], [10], [11], [12].

In the specific domains of archaeology and geomorphology, spectral indices and statistical models that condense long-term data into a limited set of products can effectively capture broad temporal changes. However, these approaches are better suited for identifying large, prominent archaeological sites (e.g. archaeological tells) and large-scale paleo-river systems [13], [14], [15] than smaller, ephemeral, or less distinct traces such as field systems, minor structures, or short-term surface modifications, which may be obscured in temporally aggregated data. At the same time, advances in quantitative methods for classifying spectral signatures of crop and soil marks – associated with vegetation stress induced by underlying archaeological and geomorphological features – have been largely confined to controlled-environment case studies [16].

Over the past five years, deep learning (DL) methods have driven significant advancements in the automated identification of objects in remote sensing imagery, primarily due to their ability to extract rich spatial and spectral features. This capability has facilitated the detection of small objects compared to the background, even under challenging conditions such as geographic diversity, varying object scales, morphologies, and sensor responses [17], [18]. In the optical domain, while substantial progress has been made in detecting surface objects, advancements in the detection and segmentation of subsoil features have lagged significantly [19], [20]. This disparity is attributable to two principal factors. Subsoil features, and more in general, all the features that rely on indirect markers, such as crop marks and soil marks, inherently lack the distinctiveness and clarity of surface elements, manifesting as subtle, low-contrast spatial and spectral patterns. These features are easily obscured by the high-frequency details of modern landscapes, such as urban areas, infrastructure, and vegetation.

Another major obstacle is the limited availability of high-quality, annotated training datasets. While surface object detection benefits from extensive and diverse datasets developed for various applications, including urban planning, security, and environmental monitoring, in contrast, subsoil detection applies to a narrower range of disciplines, such as archaeology, geology, and soil science, resulting in fewer publicly available datasets. This scarcity hinders the training and evaluation of DL models tailored for subsoil applications, impeding the replication of experiments and the comparison of findings, thereby slowing innovation and limiting the ability to benchmark novel approaches effectively [17], [21], [22], [23].

This paper introduces the MAPS dataset (Multitemporal multispectral dAtaset for Palaeochannels Segmentation), the first publicly available dataset specifically designed for the development of semantic segmentation deep neural networks aimed at identifying palaeochannels in multispectral image time series. MAPS includes annotations of palaeochannels — a specific category of subsoil features — on Sentinel-2 imagery detected at multiple temporal instances. By providing a year-long collection of labelled images at specific time intervals, the dataset enables the exploitation of a temporal series of images of the same area to maximise the likelihood of obtaining optimal visibility conditions, overcoming the challenges posed by soil and vegetation cover. By leveraging their generalisation capabilities, DL models can efficiently analyse the same area multiple times throughout the year and across several years, with minimal cost. This repeated analysis significantly increases the likelihood of detecting subtle traces that may only be visible under specific conditions. By identifying and aggregating pieces of traces visible at different times, this approach aims to reconstruct their continuity, overcoming the challenges posed by fragmented landscapes that often disrupt trace patterns and complicate their interpretation.

With the release of MAPS, we provide a benchmark of state-of-the-art semantic segmentation algorithms that can serve as a baseline for the development of future algorithms.

II. THE MAPS DATASET

A. PALAEOCHANNELS

Palaeochannels are the remnants of ancient rivers and streams that have dried up due to climatic and geological events or anthropogenic reclamation, generally no longer conveying surface water flow [24], [25]. Detecting palaeochannels contributes to understanding landscape dynamics, hydrological processes, and past land and resource management that have influenced the development of human settlements over the centuries [26], [27].

In optical imagery, paleochannels are distinguishable by the contrast between the channel fillings, composed of semi-consolidated or unconsolidated fine-grained deposits, and the coarse-grained gravelly and sandy sediments of bank and over-bank deposits. These differences lead to distinct variations in soil composition and moisture levels, which can be directly detected under bare soil conditions or by observing variations in vegetation growth.

The area analysed in this contribution extends for about 1500 Km^2 in the northern Adriatic alluvial and coastal plain across the Veneto and Friuli-Venezia Giulia regions in northern Italy (Fig. 1). This plain was formed by large fluvial systems fed by the southeastern Alps Mountain range, primarily consisting of alluvial deposits of the Last Glacial Maximum (ca. 27-17 ka cal. BP) and the late Holocene (ca. 5 ka cal. BP) [28]. The Holocene marine transgression led to the formation of lagoonal environments along the sandy barriers, most of which were reclaimed during the 20^{th} century, leading to the emersion of wide coastal areas.

B. IMAGERY COLLECTION

The dataset was created using Sentinel-2 images, which incorporate spectral bands in the 443.9 – 2202.4 nm range with variables Ground Sampling Distance of 10, 20 and 60 m.

FIGURE 1. The MAPS dataset: a) Overview of the coastal plain of the study area (northeastern Italy), divided into training, validation and test sets; b, c, d) tiles of Sentinel-2 (b4-b3-b2) images showing the changing visibility of palaeochannels on 18th April, 1st August and 24th November, respectively; e) ground-truth masks of the three images, with each colour representing a different period (April = cyan, August = orange, November = black).

The effectiveness of Sentinel-2 imagery in visually identifying palaeochannels has been demonstrated in numerous studies due to adequate spatial resolution and its multispectral and multitemporal characteristics [10], [12], [15]. Additionally, the global open-access availability of Sentinel-2, combined with its short revisiting period, allows for the analysis of large territorial areas with frequent data capture, significantly improving the replicability of experiments across various locations.

One cloud-free image per month was selected from the level-2A Sentinel-2 collection of the year 2022 (orthorectified and atmospherically corrected raster with surface reflectance values) using Google Earth Engine [29], resulting in 12 multispectral images. Since a cloud-free image for December was unavailable, it was replaced with the first one available from January 2023. The decision to compile a year-long dataset was based on the need to encompass most of the variability associated with agricultural cycle management and seasonal climate changes.

Given the gradual emergence and disappearance of palaeochannels in the time series images, the labelling process was conducted at different intervals. To identify the stages with optimal visibility of palaeochannels, we assessed 35 randomly selected palaeochannels across the study area each month of the year. Subsequently, we ranked the images based on their visibility rates and identified two primary periods: P1 = March-April and P3 = November-January. A third period (P2 = July-August) was selected to ensure comprehensive coverage throughout the year, with each period spaced two months apart (Table 1). This approach aimed to condense the yearly information into limited images while emphasising the differences between adjacent periods. The dataset area was partitioned into distinct training and validation sets, with approximately 90% allocated to training and 10% to validation (Fig. 1). The division ensured not only a spatial balance but also maintained a proportional distribution in the number of palaeochannel instances included in each subset. The percentage of tiles containing events is 89%, while those without events are 11%. This ratio remains consistent between the training and validation sets. Additionally, a geographically distinct test set was designated to assess the generalisation capabilities of the machine learning models on unseen data that represent similar environmental conditions. Covering an area of 63 km^2, this test was selected to reflect a closely related geographic and geomorphologic context (Fig. 1).

TABLE 1. The MAPS dataset structure.

	Imagery	*Labels*	*N. of instances*
Training - validation	9th March 2022	L1_train	1851
	17th April 2022		
	2nd July 2022	L2_train	1138
	1st August 2022		
	6th November 2022	L3_train	1209
	10th January 2023		
Test	18th April 2022	L1_test	126
	3rd May 2022	L2_test	70
	1st August 2022	L3_test	106
	15th October 2022	L4_test	138
	24th November 2022	L5_test	98

C. DATASET LABELS

Three operators manually labelled palaeochannels using the polygon feature class on ESRI ArcGIS [30]. The labelling

process underwent multiple stages of cross-checking and revision by all the operators to ensure consistency in the output, address any missing instances and reduce false-positive identification. Labels include various parts of the palaeochannel system, encompassing river channels, banks and over-bank deposits (see section II-A). Additionally, they incorporate buried channels of potential anthropogenic origin, distinguishable by their straight and regular pathways. The median extent of individual palaeochannel instances is 120 pixels, which, in a 10m resolution image, corresponds to an area of 12,000 m^2. The 95th percentile reaches 700 pixels, equivalent to 70,000 m^2 (Fig. 2).

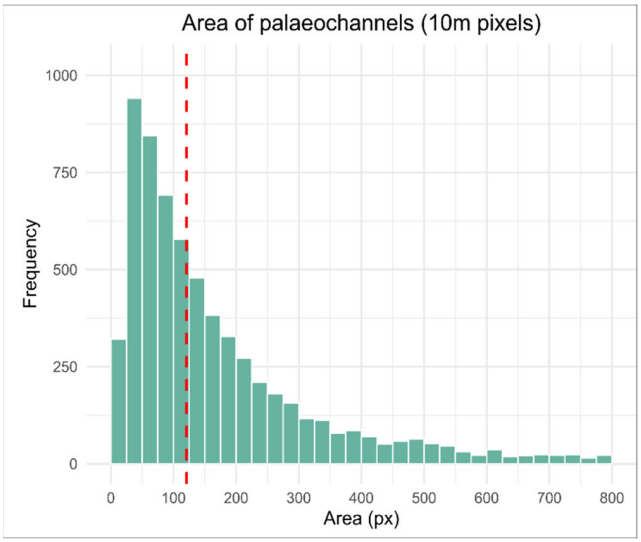

FIGURE 2. Distribution of the area of palaeochannel labels.

Three distinct groups of labels were created for each period in the training and validation area (Table 1):

- L1_train (labels on P1 images) contains 1851 instances.
- L2_train (labels on P2 images) contains 1138 instances.
- L3_train (labels on P3 images) contains 1209 instances.

In the test area, three groups of labels, namely L1_test, L2_test, and L3_test, were created based on images representing the three periods: April for P1, August for P2 and November for P3, respectively. Additionally, the test set includes labels for two months not used in training (i.e. May (L4_test) and October (L5_test)) to evaluate temporal generalisation capabilities over the year.

III. EVALUATIONS

In this section, we evaluate the performance of the three baseline semantic segmentation networks on the MAPS dataset. Segmentation, which operates at the pixel level, was selected for its ability to detect and delineate extensive and branched features like palaeochannels. This approach can accurately capture their intricate, irregular shapes and preserves the high variability in size, direction, and connectivity, maintaining the complexity of the object.

Three different types of networks have been used to assess whether the results are influenced more by the nature of the task and the data type, rather than by the specific characteristics of the model. These include two CNN-based models, U-Net [31] and YOLOv8 [32] and a Transformer-based model, UPerNet [33]. While CNNs focus on local, spatial features, Transformers are better at considering broader contextual information through self-attention, making them more effective for handling complex, context-dependent tasks.

Each architecture was trained on two different training sets, namely TR1 and TR2. In TR1, the imagery set comprises three dates, one for each period (P1, P2, P3), resulting in one image of April for P1, August for P2, and November for P3. In TR2, the imagery comprises six dates, with two images for each period, resulting in March and April for P1, July and August for P2, and November and January for P3 (Table 2).

To train the baseline models, we utilised the Red (B4), Green (B3) and Blue (B2) bands from Sentinel-2 images, as these bands were also used during the labelling process due to the high spatial resolution of 10m. As the goal of this contribution is to establish a baseline for semantic segmentation with different architectures with a clear focus on temporal assessment, the 10m NIR band (B8) and the other spectral bands at lower resolution were not utilised. However, those features are included in the dataset for further exploitation.

Images were split into 256-pixel tiles, and reflectance was normalised to the 0 to 1 range. An overlap of 50% was deliberately implemented to accommodate the extensive and intricate shapes of the labels, which frequently span adjacent tiles, allowing the models to receive comprehensive contextual information.

The resulting number of tiles is 2727 for the TR1 set and 5454 for the TR2 set (Table 2). Data augmentation methods, including vertical and horizontal flips with a 50% probability, were applied to the input data for U-Net and UPerNet. For YOLOv8, the default augmentation methods provided within the original code implementation were used, consisting of a 50% probability horizontal flip and HSV image quality adjustment with default values (HSV Hue set to 0.015, Saturation to 0.7 and Value to 0.4).

Each baseline model was trained separately on both TR1 and TR2 sets. The six trained models were then tested on five test sets, generating thirty outputs. The evaluation of segmentation results relies on the Intersection over Union (IoU), precision (P) and recall (R) (Table 3). To prioritise true positive detection and address class imbalance, the F1-score is evaluated alongside IoU, which is better suited for assessing boundary-level segmentation accuracy.

A. EXPERIMENTAL SETTING

U-Net was configured with the mit_b5 encoder, pre-trained on ImageNet. The decoder is integrated with the Spatial and Channel 'Squeeze & Excitation' (SCSE) attention module to recalibrate the learned feature maps by boosting meaningful features while suppressing the weak ones [34]. The Dice loss function was used to train the network (as well as all the other networks used in this study) as it is well-suited for highly

TABLE 2. Experimental training sets.

Imagery	*Period*	*Labels*	*N. of tiles*
TR1 set			Total: 2727 tiles
17th April 2022	P1	L1_train	909
1st August 2022	P2	L2_train	909
6th November 2022	P3	L3_train	909
TR2 set			Total: 5454 tiles
9th March 2022	P1	L1_train	909
17th April 2022			909
2nd July 2022	P2	L2_train	909
1st August 2022			909
6th November 2022	P3	L3_train	909
10th January 2023			909

unbalanced segmentation tasks [35]. The learning rate was set to 10^{-4}.

UPerNet is a general framework proposed to categorise scenes and detect objects inside an image by leveraging any vision backbone toward the design of a vision Transformer [36]. UPerNet framework for semantic segmentation consists of several components, including a visual backbone, a Feature Pyramid Network (FPN) and a Pyramid Pooling Module (PPM). In this study, UPerNet was configured with the ConvNeXt backbone, pre-trained on ImageNet and ADE20k datasets. ConvNeXt represents an evolution of the standard ResNet, adopting a design akin to the hierarchical Transformer Swin, while retaining the simplicity characteristics of a standard ConvNet [37]. It demonstrated competitive performance across different model capacities. A learning rate of 10^{-4} was used to train the network.

YOLOv8 for segmentation, a variant of the YOLO series, was used for this study. The YOLOv8 is pre-trained on the COCO dataset and provides five models differing in size and number of parameters. After experimenting with these models, the ''yolov8n-seg'' lightweight model was selected because of its performance and training speed. A learning rate of 10^{-3} was used to train the network.

The details about the hyperparameters and the number of epochs are included in the Hugging Face and IIT Dataverse repositories.

To evaluate the robustness of the models across varying conditions, we conducted four training runs with different random seeds. To determine whether the observed differences in performance were statistically significant, we performed an Analysis of Variance (ANOVA) of IoU scores. Pairwise comparisons were then conducted to identify specific group differences, specifically to assess the significance of the results using the two distinct training sets (TR1 vs TR2) and segmentation architecture comparison.

B. SEASONAL CONTRIBUTION

Beyond assessing general performance, this study aims to evaluate the contribution of each observation (monthly image) to the reconstruction of the full extent of the palaeochannel traces. The underlying premise is that data collected at different times of the year provide complementary insights, thereby improving the likelihood of detecting short-lived traces that appear only under specific environmental conditions. Detecting these traces greatly contributes to the overall identification of palaeochannels and addresses the visibility challenges outlined in the introduction. Each image's contribution is quantified by counting the number of newly identified pixels that do not overlap with previously predicted labels from earlier months. This method determines the extent to which each observation provides new information for delineating the actual shape of the palaeochannel traces.

To measure this contribution, a chronological order was followed, treating the dataset as a one-year period. Beginning with April (the first month of the test dataset), the additional contribution of each subsequent month was assessed in terms of newly detected pixels. Only true positive (TP) predictions were considered in this analysis.

C. RESULTS

The results (Tables 3, 4, 5) reveal that TR2 outperforms TR1 under some specific configurations. For example, U-Net trained with TR2 achieved a higher mean IoU of 0.23 in April, compared to 0.21 for TR1, with statistical significance ($p < 0.05$). Similarly, UPerNet trained with TR2 showed improved performance in November (0.28) and October (0.14) compared to TR1 (0.26, 0.13), with statistical significance ($p < 0.01$, $p < 0.05$). However, the large standard deviations relative to the means and the limited variability in performance indicate significant fluctuations in the results, making it difficult to draw a definitive conclusion about which training configuration is consistently superior across all models and images.

When comparing architectures, UPerNet and U-Net consistently outperformed YOLOv8 across both training sets (TR1 and TR2) over the different test set images. The statistical analysis confirmed significant differences between UPerNet and U-Net, with UPerNet achieving superior performance in April and November with TR1 (0.24, 0.25, respectively) and in April with TR2 (0.25) ($p < 0.05$). These results position UPerNet as the best-performing model under specific conditions.

Additionally, a trade-off between precision and recall was observed. UPerNet and U-Net generally exhibited higher recall, with their best values recorded in November under TR1 (0.46 and 0.52, respectively), significantly outperforming YOLOv8's highest recall value of 0.23. Conversely, YOLOv8 demonstrated higher precision, peaking at 0.50 in April under TR1. However, its performance was inconsistent due to high standard deviations compared to U-Net and UPerNet.

Considering seasonal variations, the highest IoU was achieved in November using TR2, with UPerNet reaching 0.28 and U-Net following closely at 0.26. Using TR1, U-Net recorded 0.27 in November, while UPerNet reached

TABLE 3. Results of the TR1 models.

Months	U-Net				UPerNet				YOLOv8			
	TR1				TR1				TR1			
	IoU	Precision	Recall	F1	IoU	Precision	Recall	F1	IoU	Precision	Recall	F1
April	0.21 ±0.01	0.32 ±0.06	0.43 ±0.04	0.36 ±0.03	0.24 ±0,01	0.39 ±0.01	0.40 ±0.03	0.40 ±0.02	0.14 ±0.01	0.5 ±0.03	0.16 ±0.02	0.24 ±0.02
May	0.19 ±0.01	0.33 ±0.06	0.36 ±0.04	0.34 ±0.02	0.22 ±0.02	0.34 ±0.09	0.33 ±0.05	0.32 ±0.04	0.13 ±0.02	0.3 ±0.11	0.20 ±0.01	0.24 ±0.03
August	0.08 ±0.02	0.21 ±0.03	0.12 ±0.04	0.15 ±0.03	0.10 ±0.03	0.25 ±0.02	0.15 ±0.05	0.18 ±0.04	0.04 ±0.01	0.14 ±0.02	0.07 ±0.02	0.09 ±0.02
October	0.14 ±0.01	0.31 ±0.02	0.22 ±0.05	0.25 ±0.03	0.13 ±0.01	0.40 ±0.04	0.17 ±0.02	0.24 ±0.02	0.10 ±0.09	0.26 ±0.26	0.07 ±0.02	0.10 ±0.04
November	***0.27 ±0.01***	0.37 ±0.04	0.52 ±0.05	0.43 ±0.01	***0.26 ±0.01***	0.38 ±0.04	0.46 ±0.04	0.41±0.01	***0.15 ±0.04***	0.35 ±0.17	0.23 ±0.02	0.27 ±0.05

IoU, precision, recall and F1 for all the outputs from the training set TR1. In bold-italic: for each architecture, the best-performing monthly image (IoU). In underline: for each monthly image, the best-performing architecture (IoU)

TABLE 4. Results of the TR2 models.

Months	U-Net				UPerNet				YOLOv8			
	TR2				TR2				TR2			
	IoU	Precision	Recall	F1	IoU	Precision	Recall	F1	IoU	Precision	Recall	F1
April	0.23 ±0.01	0.37 ±0.03	0.43± 0.05	0.40± 0.02	0.25 ±0.00	0.37 ±0.02	0.45 ±0.02	0.40 ±0.01	0.15 ±0.02	0.36 ±0.06	0.21 ±0.01	0.27 ±0.03
May	0.19 ±0.03	0.34 ±0.13	0.38 ±0.12	0.33± 0.04	0.20 ±0.01	0.37 ±0.02	0.30 ±0.01	0.33 ±0.01	0.12 ±0.05	0.34 ±0.09	0.16 ±0.05	0.22 ±0.06
August	0.09 ±0.03	0.20 ±0.09	0.15 ±0.08	0.15 ±0.05	0.07 ±0.04	0.26 ±0.02	0.10 ±0.07	0.14 ±0.07	0.05 ±0.03	0.23 ±0.17	0.07 ±0.04	0.10 ±0.06
October	0.13 ±0.03	0.34 ±0.11	0.21 ±0.10	0.23 ±0.04	0.14 ±0.01	0.44 ±0.06	0.18 ±0.01	0.25 ±0.01	0.04 ±0.01	0.21 ±0.10	0.05 ±0.02	0.08 ±0.03
November	***0.26 ±0.04***	0.41 ±0.13	0.48 ±0.11	0.42 ±0.06	***0.28 ±0.01***	0.42 ±0.01	0.46 ±0.03	0.44 ±0.02	***0.16 ±0.03***	0.42 ±0.08	0.22 ±0.02	0.28 ±0.03

IoU, precision, recall and F1 for all the outputs from the training set TR2. In italic: best IoU per network across the monthly images. In underline: best IoU per monthly image across different networks

0.26. YOLOv8's best IoU performance was significantly lower, peaking at 0.16 with TR2 in November. In descending order of IoU performance, the best-performing months were November, April (0.25 with TR2 UPerNet), May (0.20 with TR2 UPerNet and U-Net), October (0.14 with TR2 UPerNet), and August (0.10 with TR1 UPerNet). This trend is largely consistent across all configurations.

The data generally supports a positive correlation between IoU and F1 scores, suggesting that improvements in spatial overlaps often coincide with improvements in overall performance. Although minor differences are present between the two metrics in specific instances, these do not fundamentally affect the primary findings of this analysis.

IV. DISCUSSION

An increase in dataset size from TR1 to TR2 resulted in a slight improvement in model performance. However, this improvement was statistically significant only for specific configurations, including U-Net in April and UPerNet in October and November (Table 5). Despite TR2 being twice the size of TR1, the additional data used to train the TR2 models may not have provided as much relevant information as initially anticipated. This outcome is likely due to the TR2 dataset characteristics, which increase the number of temporal observations by including two consecutive images for each period, separated by approximately one month, and using the same set of labels (as described in Section III). While the subtle changes in land cover and vegetation between these images allowed for consistent labelling, they were not diverse enough to provide significant new information for the models. The limited variation between these consecutive images restricted the models' ability to fully use the additional data, preventing substantial performance improvements despite the larger dataset size.

TABLE 5. Results of pairwise ANOVA.

Model Comparison	Month	F-Value	p-Value	Statistical Significance
Networks				
TR1 U-Net vs. UPerNet				
	April	13.44	0.01	**
	November	6.00	0.04	**
TR2 U-Net vs. UPerNet				
	April	8.16	0.03	**
Training sets				
U-Net TR1 vs. TR2				
	April	12.78	0.01	**
UPerNet TR1 vs. TR2				
	November	24.20	0.002	***
	October	7.19	0.03	**

Pairwise ANOVA ($\alpha = 0.05$) assessing significant performance differences between models (U-Net, UPerNet) and training sets (TR1, TR2). Only statistically significant results are shown. Significance levels: ** $p < 0.05$, *** $p < 0.01$.

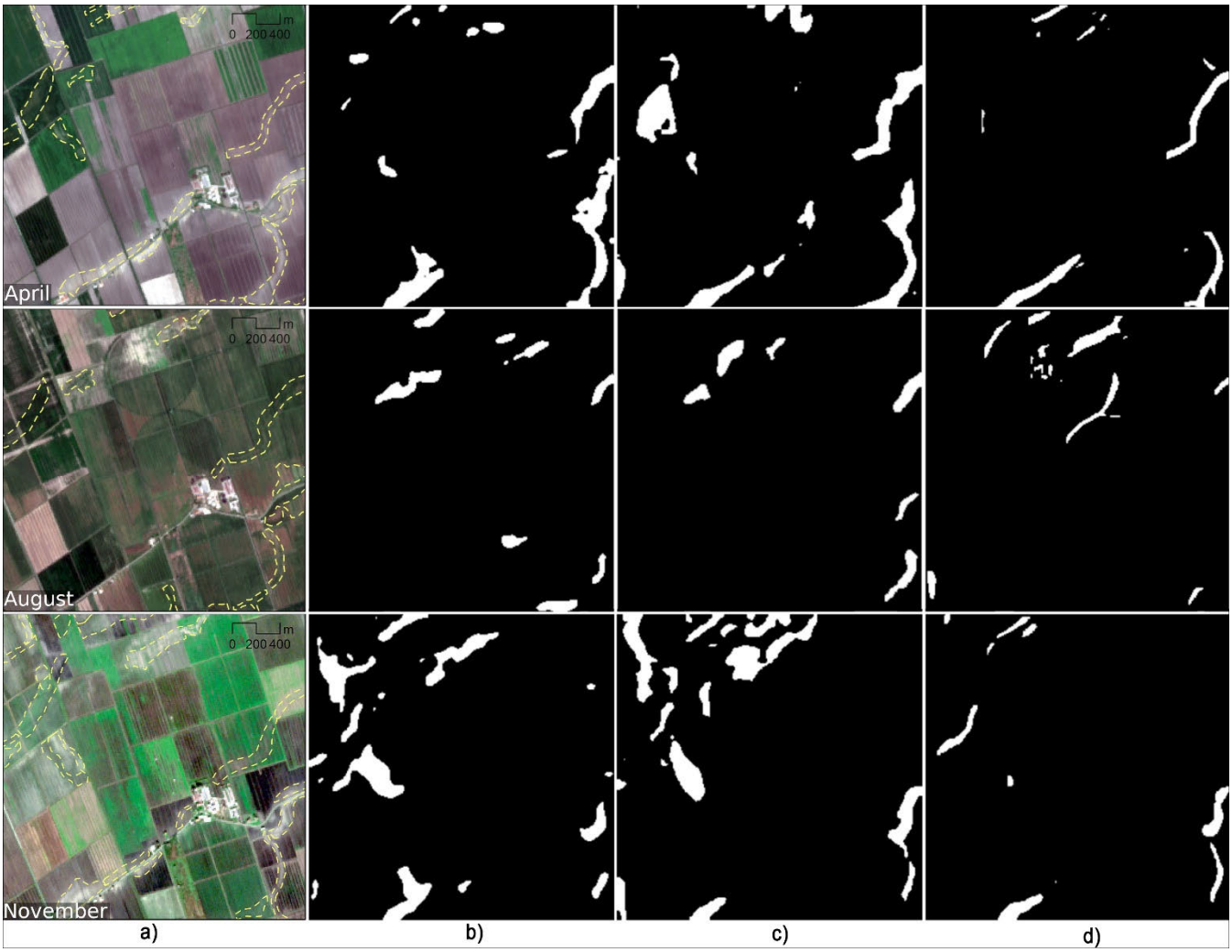

FIGURE 3. Visualisation of the results on April, August and November images: a) Sentinel-2 image with overlapping ground-truth masks (dotted yellow lines); b) U-Net predicted masks; c) UPerNet predicted masks; d) YOLOv8 predicted masks.

All models consistently demonstrated similar performance patterns across the monthly images. The highest Intersection over Union (IoU) scores were achieved for the November image, followed by April and May. Conversely, August and October showed the lowest IoU values. A visual inspection of the test images from August revealed increased vegetation coverage during this month, likely due to the peak of agricultural activity (Fig. 3). Despite being one of the driest months of the year, artificial irrigation helped maintain healthy vegetation with minimal water stress. Vegetation traces are typically good indicators for detecting subsoil archaeological and geomorphological features to human observers [3], [5], [6]. However, in this case, dense, healthy vegetation makes it more difficult to identify palaeochannel traces. The well-irrigated conditions reduce vegetation stress, which in turn diminishes the visibility of crop marks associated with subtle subsoil features. This challenge contrasts with the high number of palaeochannels identified by manual operators (see Table 1 for instance counts). The models appear to face greater difficulties than human observers, likely due to the subtlety of features that require specialised attention in deep learning development. Additionally, artificial irrigation alters the uniformity of the field surface, creating moisture patterns (e.g., circular marks in Fig. 3) that further obscure palaeochannel visibility.

Similarly, low performance was observed in October, despite it being the month with the highest number of manually detected instances. Image inspection shows that the harvesting had concluded in most fields, resulting in exposure to bare soil conditions. However, the dry conditions during this period hindered the distinct appearance of soil marks, as the reduced moisture caused the traces to become blurred and hard to distinguish for the model.

Conversely, the combination of bare soil and high moisture content, particularly in April, November and May, created more favourable conditions for the automated models. The richer contrasts in these images enhanced their ability to

detect palaeochannels, thereby improving detection accuracy (Fig. 3).

A dedicated evaluation of May and October, the two months excluded from the training set, provides insight into the models' generalisation performance. These months exhibit contrasting results, with higher performance in May and lower performance in October. Both are characterised by low vegetation and predominant bare soil conditions; however, May shows more vegetation and higher moisture content, with fewer manually detected instances, while October shows drier conditions and the highest number of instances.

Performance in May is comparable to that of April and November, though slightly lower, likely due to the reduced number of visible instances that indicate less favourable conditions. In contrast, the lower performance in October appears to result from the challenging environmental conditions rather than a limitation in the model's ability to generalise.

In Fig. 3, the visual assessment of predicted segmentation masks from models trained on TR2 provides a comprehensive means of interpreting quantitative precision and recall values, gaining deeper insights into the comparative performance of the models. UPerNet and U-Net, which exhibit higher recall values, result in larger and less well-defined masks that often extend beyond the ground-truth boundaries. While this diffusion and extension successfully reduce the number of undetected features, it negatively impacts precision, leading to lower values. Conversely, YOLOv8 produces masks that more precisely overlap the ground truth labels, achieving higher precision. However, this comes at the cost of a high percentage of false negative instances, meaning that many instances remain undetected.

The segmentation labels generated by all models contribute to reconstructing the intricate shapes of the palaeochannels. While the accuracy may be compromised, the segmentation masks preserve sufficient structural detail to capture the complex, meandering forms of the palaeochannels, making them valuable for further analysis and interpretation in prospection activities.

A. SEASONAL CONTRIBUTION

The analysis of newly detected pixels across multiple months provides valuable insights into how different time periods contribute to expanding the mapped features of palaeochannels. Figure 4 illustrates the monthly trend of newly detected pixels, spanning from April through the entire year. The results show that May and November contributed most significantly to expanding April's labels, each accounting for 14% of the total true positive (TP) pixels. August followed with an 8% contribution, while October added only 4%. May showed significant potential despite sharing similar environmental conditions with April and exhibiting the lowest overall number of instances (Tab. 1). The rapid changes in vegetation growth and soil moisture content typical of this period of

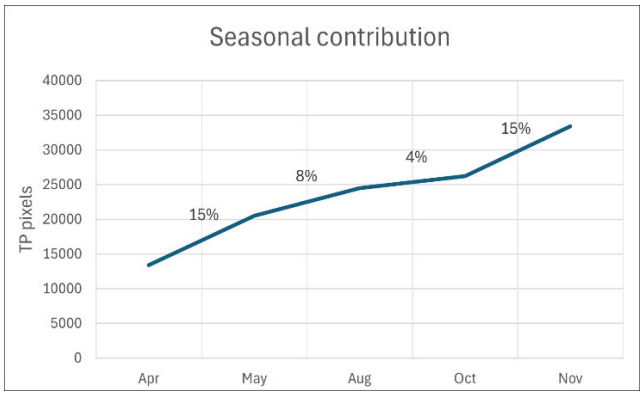

FIGURE 4. Contribution of different images analysed throughout the year in chronological order, relative to the first and most contributing image (April). The steepness of the line represents the absolute contribution to expanding predicted labels with new, non-overlapping labels. Only true positive (TP) pixels are considered.

the year likely created altered field conditions, leading to the emergence of new crop marks. Additionally, variations in crop types and field rotations likely influenced visibility, enabling the detection of new traces to appear. November represents a post-harvest period with largely bare soil conditions and high moisture content, similar to April. However, the significant contribution to detecting new features suggests that it provides a complementary perspective, likely due to agricultural cycles and crop rotation. In August, the presence of dense green vegetation in cultivated fields reveals crop marks that differ from those observed in bare soil. The contrasting seasonal conditions further contribute to the appearance of new features, accounting for more than 80% of the total pixels identified in the image. However, October yielded the lowest performance, both in absolute terms and in relative increments of new features. This suggests that the features detected in October had already been visible in previous months, despite the unique dry conditions and the fact that many agricultural fields were left uncultivated. As a result, October's characteristics largely mirrored those of the preceding months, providing a minimal contribution to the expansion of the palaeochannel traces map.

The map of newly detected true positive (TP) pixels across various months (Figure 5) illustrates how this approach can effectively fill gaps caused by fragmentation from variable agricultural field patterns, seasonal conditions, and agricultural activities. The accumulation of TP pixels over time creates a more continuous map of palaeochannels, enhancing overall coverage and reducing the impact of missing or fragmented features in individual observations. By incorporating multiple temporal perspectives, patterns and structural details are revealed that may not be visible in a single acquisition. Changes in agricultural fields, such as planting, harvesting, and crop rotation, enhance the robustness of palaeochannels detection. This temporal complementarity improves the identification and visibility of subtle patterns, which are crucial for the interpretation of the landscape's historical characteristics.

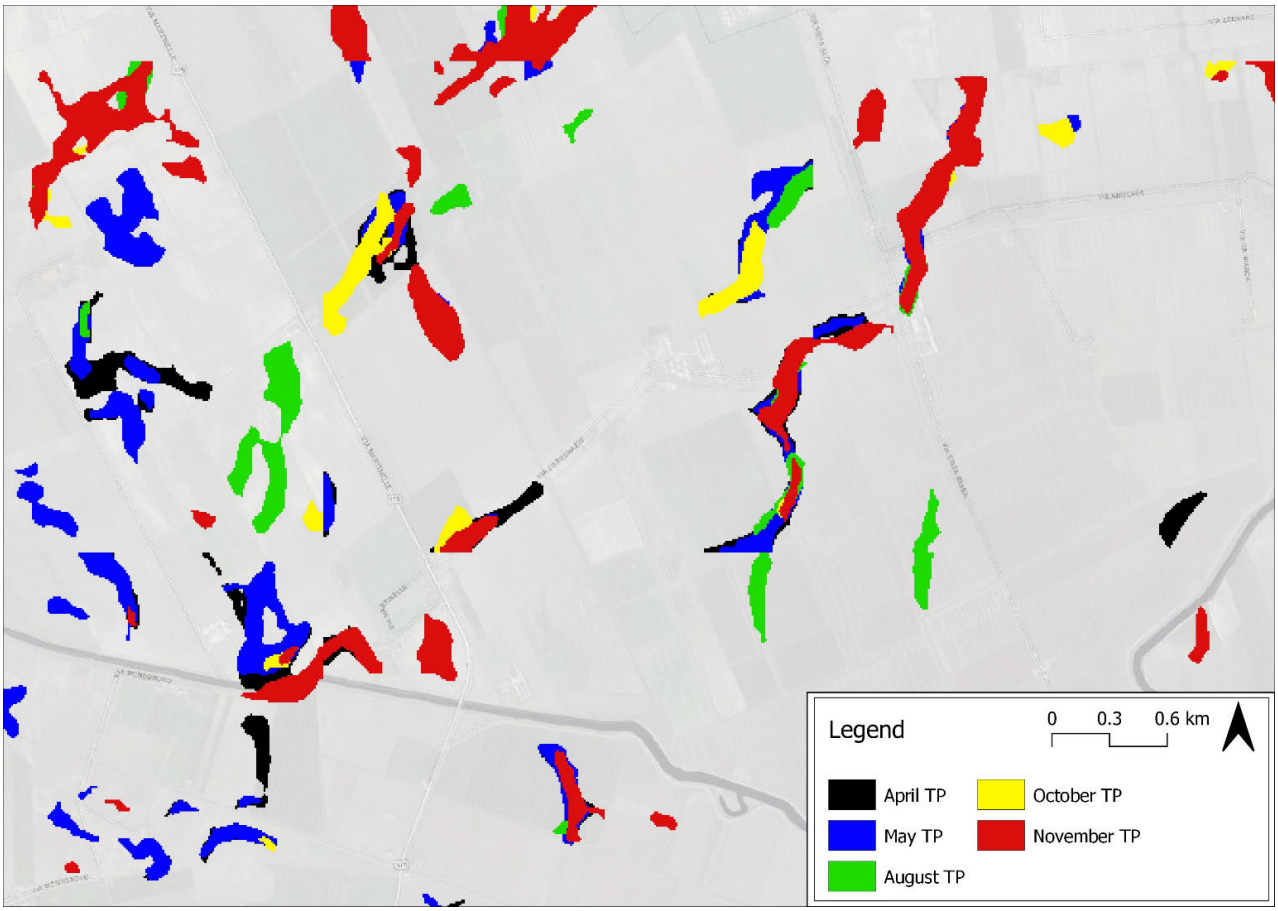

FIGURE 5. The plot of true positive (TP) pixel labels from inference on all test images enables visual assessment of the seasonal contribution of each image in clarifying and expanding the continuity of palaeochannel traces. Sample shown from the test area.

V. CONCLUSION

The MAPS dataset introduced here has demonstrated its ability to ensure the reproducibility of results and to provide a valuable foundation for advancing automated mapping techniques. Tests conducted with the dataset further underscore the importance of incorporating temporal series analysis and carefully designed architectural configurations to enhance segmentation accuracy. Temporal series analysis, in particular, enables a deeper understanding of changes and patterns over time, offering valuable insights into optimal visibility conditions of markers in soil and vegetation, insights that static data alone cannot provide. Results indicate that variability in contiguous images can be as relevant as the differences observed across distinct periods of the year, such as those observed between April and May in our case study. This variability is driven by short-term factors and modifications, including climate fluctuations, changes in precipitation levels, variations in soil moisture, and agricultural practices like crop rotation and harvesting schedules. Such variations can only be effectively captured through repeated observations, underlying the importance of temporal series analysis. The analysis of seasonal contributions confirms that even images with apparent lower visibility (e.g. August) should not be dismissed, as they capture unique vegetation conditions and reflect variations in agricultural cycles. Therefore, testing multiple observations can significantly enhance the spatial continuity of traces, leading to a more comprehensive mapping of palaeochannels and improving the interpretability of traces that might otherwise appear fragmented or only partially visible.

Additionally, the choice and design of architectural configurations influence both inference quality and the accuracy of predicted masks. These should be evaluated not only by IoU scores but also by their practical utility in archaeological and geomorphological prospection. Even low-performing segmentation models can delineate complex subsoil features such as palaeochannels, whose elongated and ramified structures resemble biological formations rather than standard objects found in commonly used training datasets. This segmentation provides information on orientation, direction, continuity, and relationships with other traces, making it particularly useful for interpreting landscapes and their historical features. Incorporating data from multiple time periods enhances the identification of palaeochannel extent, enabling a more comprehensive understanding of the landscape. This approach supports environmental monitoring, land man-

agement, and conservation, where accurate knowledge of palaeochannels and geomorphological features informs sustainable practices and long-term landscape stewardship.

Future research should focus on improving model performance during months with lower segmentation accuracy, which may be due to architectural limitations or suboptimal data conditions. Exploring advanced spatiotemporal attention mechanisms or transformer-based hybrid architectures could enhance generalisation and capture complex temporal patterns. The integration of multispectral bands into the MAPS dataset enables the development of data augmentation techniques, such as spectral index extraction to reduce vegetation interference, spectral unmixing, and dimensionality reduction, which could mitigate environmental challenges like dense vegetation or moisture variations. Systematic experimentation with additional spectral bands, including SWIR bands (e.g., Sentinel-2 B11 and B12), despite their lower spatial resolution, may further improve subsoil feature detection. Integrating complementary remote sensing data, such as Synthetic Aperture Radar or airborne LiDAR, could also enhance the identification of subtle topographic, soil, or crop-related markers.

Subsequent research efforts could leverage MAPS to validate and expand these methodologies across diverse geographical contexts and environmental conditions, exploring their effectiveness in detecting a broader range of subsoil features, including archaeological structures, ancient roads, and geomorphological traces beyond palaeochannels. Addressing these challenges has the potential to improve current segmentation approaches and foster further interdisciplinary research in landscape analysis.

ACKNOWLEDGMENT

(Giulio Poggi and Andaleeb Yaseen contributed equally to this work.)

REFERENCES

[1] R. Lasaponara and N. Masini, "Satellite remote sensing in archaeology: Past, present and future perspectives," *J. Archaeological Sci.*, vol. 38, no. 9, pp. 1995–2002, Sep. 2011, doi: 10.1016/j.jas.2011.02.002.

[2] W. A. Marcus and M. A. Fonstad, "Remote sensing of rivers: The emergence of a subdiscipline in the river sciences," *Earth Surf. Processes Landforms*, vol. 35, no. 15, pp. 1867–1872, Dec. 2010, doi: 10.1002/esp.2094.

[3] A. Agapiou, D. Hadjimitsis, and D. Alexakis, "Evaluation of broadband and narrowband vegetation indices for the identification of archaeological crop marks," *Remote Sens.*, vol. 4, no. 12, pp. 3892–3919, Dec. 2012, doi: 10.3390/rs4123892.

[4] G. J. Verhoeven, "Near-infrared aerial crop mark archaeology: From its historical use to current digital implementations," *J. Archaeological Method Theory*, vol. 19, no. 1, pp. 132–160, Mar. 2012, doi: 10.1007/s10816-011-9104-5.

[5] N. Masini, C. Marzo, P. Manzari, A. Belmonte, C. Sabia, and R. Lasaponara, "On the characterization of temporal and spatial patterns of archaeological crop-marks," *J. Cultural Heritage*, vol. 32, pp. 124–132, Jul. 2018, doi: 10.1016/j.culher.2017.12.009.

[6] G. Verhoeven, "Are we there yet? A review and assessment of archaeological passive airborne optical imaging approaches in the light of landscape archaeology," *Geosciences*, vol. 7, no. 3, p. 86, Sep. 2017, doi: 10.3390/geosciences7030086.

[7] A. Agapiou, D. G. Hadjimitsis, A. Sarris, A. Georgopoulos, and D. D. Alexakis, "Optimum temporal and spectral window for monitoring crop marks over archaeological remains in the Mediterranean region," *J. Archaeological Sci.*, vol. 40, no. 3, pp. 1479–1492, Mar. 2013, doi: 10.1016/j.jas.2012.10.036.

[8] S. Peña-Villasenín, M. Gil-Docampo, and J. Ortiz-Sanz, "Hidden archaeological remains in heterogeneous vegetation: A crop marks study in fortified settlements of Northwestern Iberian Peninsula," *Remote Sens.*, vol. 16, no. 21, p. 3923, Oct. 2024, doi: 10.3390/rs16213923.

[9] R. Lasaponara, N. Abate, and N. Masini, "On the use of Google Earth engine and sentinel data to detect 'lost' sections of ancient roads. The case of via appia," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2021.3054168.

[10] F. Brandolini, G. Domingo-Ribas, A. Zerboni, and S. Turner, "A Google Earth engine-enabled Python approach for the identification of anthropogenic Palaeo-landscape features," *Open Res. Eur.*, vol. 1, p. 22, Sep. 2021, doi: 10.12688/OPENRESEUROPE.13135.2.

[11] I. D. Negula, C. Moise, A. M. Lazăr, N. C. Rişcuţa, C. Cristescu, A. L. Dedulescu, C. E. Mihalache, and A. Badea, "Satellite remote sensing for the analysis of the micia and germisara archaeological sites," *Remote Sens.*, vol. 12, no. 12, p. 2003, Jun. 2020, doi: 10.3390/rs12122003.

[12] N. Abate, A. Elfadaly, N. Masini, and R. Lasaponara, "Multitemporal 2016–2018 Sentinel-2 data enhancement for landscape archaeology: The case study of the foggia province, Southern Italy," *Remote Sens.*, vol. 12, no. 8, p. 1309, Apr. 2020, doi: 10.3390/rs12081309.

[13] R. Valente, E. Maset, and M. Iamoni, "Three years of Google Earth engine-based archaeological surveys in Iraqi kurdistan: Results from the ground," *Remote Sens.*, vol. 16, no. 22, p. 4229, Nov. 2024, doi: 10.3390/rs16224229.

[14] A. Agapiou, "Remote sensing heritage in a petabyte-scale: Satellite data and heritage Earth engine applications," *Int. J. Digit. Earth*, vol. 10, no. 1, pp. 85–102, Jan. 2017, doi: 10.1080/17538947.2016.1250829.

[15] H. Orengo and C. Petrie, "Large-scale, multi-temporal remote sensing of Palaeo-river networks: A case study from Northwest India and its implications for the Indus civilisation," *Remote Sens.*, vol. 9, no. 7, p. 735, Jul. 2017, doi: 10.3390/rs9070735.

[16] A. Agapiou and E. Gravanis, "A machine-learning-assisted classification algorithm for the detection of archaeological proxies (cropmarks) based on reflectance signatures," *Remote Sens.*, vol. 16, no. 10, p. 1705, May 2024, doi: 10.3390/rs16101705.

[17] Y. Wang, S. M. A. Bashir, M. Khan, Q. Ullah, R. Wang, Y. Song, Z. Guo, and Y. Niu, "Remote sensing image super-resolution and object detection: Benchmark and state of the art," *Expert Syst. Appl.*, vol. 197, Jul. 2022, Art. no. 116793, doi: 10.1016/j.eswa.2022.116793.

[18] J. Li, Y. Cai, Q. Li, M. Kou, and T. Zhang, "A review of remote sensing image segmentation by deep learning methods," *Int. J. Digit. Earth*, vol. 17, no. 1, Dec. 2024, Art. no. 2328827, doi: 10.1080/17538947.2024.2328827.

[19] S. Gui, S. Song, R. Qin, and Y. Tang, "Remote sensing object detection in the deep learning era—A review," *Remote Sens.*, vol. 16, no. 2, p. 327, Jan. 2024, doi: 10.3390/rs16020327.

[20] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 166–177, Jun. 2019, doi: 10.1016/j.isprsjprs.2019.04.015.

[21] I. Kadhim and F. M. Abed, "A critical review of remote sensing approaches and deep learning techniques in archaeology," *Sensors*, vol. 23, no. 6, p. 2918, Mar. 2023, doi: 10.3390/s23062918.

[22] J. Lv, Q. Shen, M. Lv, Y. Li, L. Shi, and P. Zhang, "Deep learning-based semantic segmentation of remote sensing images: A review," *Frontiers Ecology Evol.*, vol. 11, Jul. 2023, Art. no. 1201125, doi: 10.3389/fevo.2023.1201125.

[23] M. Fiorucci, M. Khoroshiltseva, M. Pontil, A. Traviglia, A. Del Bue, and S. James, "Machine learning for cultural heritage: A survey," *Pattern Recognit. Lett.*, vol. 133, pp. 102–108, May 2020, doi: 10.1016/j.patrec.2020.02.017.

[24] R. A. L. Wray, "Palaeochannels of the namoi river floodplain, new south wales, Australia: The use of multispectral Landsat imagery to highlight a late quaternary change in fluvial regime," *Austral. Geographer*, vol. 40, no. 1, pp. 29–49, Mar. 2009, doi: 10.1080/00049180802656952.

[25] M. Abdulkarim, S. Chapkanski, D. Ertlen, H. Mahmood, E. Obioha, F. Preusser, C. Rambeau, F. Salomon, M. Schiemann, and L. Schmitt, "Morpho-sedimentary characteristics of Holocene paleochannels in the upper Rhine alluvial plain, France," *E&G Quaternary Sci. J.*, vol. 71, no. 2, pp. 191–212, Sep. 2022, doi: 10.5194/egqsj-71-191-2022.

[26] M. Bini, G. Zanchetta, E. Regattieri, I. Isola, R. N. Drysdale, F. Fabiani, S. Genovesi, and J. C. Hellstrom, "Hydrological changes during the Roman climatic optimum in Northern Tuscany (central Italy) as evidenced by speleothem records and archaeological data," *J. Quaternary Sci.*, vol. 35, no. 6, pp. 791–802, Aug. 2020, doi: 10.1002/jqs.3224.

[27] P. Mozzi, S. Piovan, and E. Corró, "Long-term drivers and impacts of abrupt river changes in managed lowlands of the adige river and Northern po delta (Northern Italy)," *Quaternary Int.*, vol. 538, pp. 80–93, Feb. 2020, doi: 10.1016/j.quaint.2018.10.024.
[28] A. Fontana, "Introduction to the thematic issue: 'Alluvial geomorphology in Italy,'" *Géomorphologie Relief Process. Environ.*, vol. 18, no. 2, pp. 123–130, Nov. 2012, doi: 10.4000/geomorphologie.9792.
[29] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, Dec. 2017, doi: 10.1016/j.rse.2017.06.031.
[30] *ESRI*. Accessed: Mar. 7, 2025. [Online]. Available: http://www.esri.com
[31] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," 2015, *arXiv:1505.04597*.
[32] D. Reis, J. Kupec, J. Hong, and A. Daoudi, "Real-time flying object detection with YOLOv8," 2023, *arXiv:2305.09972*.
[33] T. Xiao, Y. Liu, B. Zhou, Y. Jiang, and J. Sun, "Unified perceptual parsing for scene understanding," 2018, *arXiv:1807.10221*.
[34] A. G. Roy, N. Navab, and C. Wachinger, "Concurrent spatial and channel squeeze & excitation in fully convolutional networks," 2018, *arXiv:1803.02579*.
[35] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571, doi: 10.1109/3DV.2016.79.
[36] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," 2022, *arXiv:2201.03545*.
[37] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 9992–10002, doi: 10.1109/ICCV48922.2021.00986.

GIULIO POGGI received the B.A. and M.A. degrees in archaeology and the Ph.D. degree in environmental, geological, and polar sciences and technologies from the University of Siena, Italy, in 2021. He held a visiting research position at the University of Lund, Sweden, in 2015. Since 2023, he has been a Postdoctoral Researcher at the Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, Italy. He works on exploring the potential of automated machine learning methods in the analysis of diverse remote sensing data for archaeological prospection. His research interests include landscape archaeology and the use of remote sensing technologies and geospatial analysis for archaeological studies.

ANDALEEB YASEEN received the B.S. degree in environmental sciences from Fatima Jinnah Women's University, Rawalpindi, Pakistan, and the M.S. degree from International Islamic University, Islamabad, Pakistan. She is currently pursuing the joint Ph.D. degree with Ca' Foscari University and the Istituto Italiano di Tecnologia, Italy. She studied at the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University, China. Her research interests include the detection of subsoil natural and anthropogenically modified features using advanced machine learning models on remote sensing data.

RAVEERAT JATURAPITPORNCHAI received the B.S. degree in computer engineering from the King Mongkut's Institute of Technology, Bangkok, Thailand, in 2013, the M.S. degree in information and communication technology for embedded systems from Kasetsart University, Bangkok, in 2016, and the Ph.D. degree in architecture and building engineering from Tokyo Institute of Technology, Kanagawa, Japan, in 2020. She was a Postdoctoral Researcher at the Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, Italy, from 2023 to 2024. Her research interests include machine learning, computer vision, and remote sensing.

SARA FERRO received the M.S. degree in control and automation engineering from the University of Padua, Italy, in 2017, and the Ph.D. degree in computer science from the Ca' Foscari University of Venice, Italy, in 2024. From 2017 to 2020, she was a Computer Vision Engineer, developing software for defect detection on cables. Currently, she is a Postdoctoral Researcher with the Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, Italy. She has been developing models for automatic historical handwritten text recognition (HTR), aiming at incorporating contextual information. Her research interests include natural language processing and computer vision.

GREGORY SECH received the B.S. degree in computer science from the Ca' Foscari University of Venice, in 2019. From 2022 to 2024, he was a Research Fellow of machine learning at the Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, Venice, Italy. He has contributed to developing projects aimed at automatically detecting looting activities and sub-soil archaeological features on satellite imagery with computer vision approaches. His research interests include machine learning applications for remote sensing data and hyperspectral image processing.

PETER NAYLOR received the Engineering Diploma degree in data science from ENSAE Paris, the master's degree in statistics from Paris Dauphine University, and the Ph.D. degree in bioinformatics from Mines Paris—PSL, with a focus on breast cancer treatment response based on histopathology slides.

He was a Postdoctoral Fellow with the High-Dimensional Statistical Modeling Team, RIKEN-AIP, Japan. He is currently a Postdoctoral Fellow with the Φ-Laboratory, ESA, Italy. His research interests include computer vision and its applications to earth observation, as well as explainable and interpretable modeling.

MARIA CRISTINA SALVI received the M.S. degree in archaeology from the University of Florence, Italy, in 2005, the master's degree in geo-technologies for archaeology from the University of Siena, Italy, in 2007, and the Ph.D. degree in sciences and technologies for archaeology and cultural heritage from the University of Ferrara, Italy, in 2012. She was an Assistant Researcher with the Center of Geotechnologies, University of Siena, and a Researcher for private companies. From 2022 to 2023, she was a Postdoctoral Researcher at the Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, Italy. Currently, she is a Researcher with the Eratosthenes Center of Excellence, Cyprus. Her research interests include the use of remote sensing technologies for the analysis of cultural heritage sites.

SEBASTIANO VASCON received the Ph.D. degree from the Istituto Italiano di Tecnologia, Genoa, Italy, in 2016, under the supervision of Prof. Vittorio Murino. In 2017, he moved to the Ca' Foscari University of Venice, Italy, as a Postdoctoral Researcher of machine learning within the Group of Prof. Marcello Pelillo. He is currently an Assistant Professor with the Department of Environmental Sciences, Computer Science and Statistics (DAIS), Ca' Foscari University of Venice. His research interests include computer vision, game theory, machine learning, and deep learning.

BERTRAND LE SAUX (Senior Member, IEEE) received the M.S.Eng. and M.Sc. degrees in signal processing from the Institut Polytechnique de Grenoble, Grenoble, France, in 1999, the Ph.D. degree in computer science from the University of Versailles/Inria, Versailles, France, in 2003, and the Habilitation degree in physics from the University of Paris-Saclay, Saclay, France, in 2019. In 2024, he founded his own consulting company, AI4Earth. His research interests include the visual understanding of the environment by data-driven techniques, including computer vision and (quantum) machine learning. He is also interested in tackling practical problems that arise in Earth observation, in bringing solutions to the current environmental and societal challenges. He was the Co-Chair of the IEEE GRSS Technical Committee on Image Analysis and Data Fusion, from 2015 to 2017, where he was the Chair, from 2017 to 2019. He is an Associate Editor of IEEE Geoscience and Remote Sensing Letters. He co-organizes many workshops and events in machine learning for Earth observation, notably the CVPR Earth Vision Workshop Series and the ESA-CMWF Workshop Series on Machine Learning for Earth System Observation and Prediction.

MARCO FIORUCCI received the M.S. and Ph.D. degrees in computer science from the Ca' Foscari University of Venice, Italy, in 2015 and 2019, respectively. Since 2019, he has been a Researcher of machine learning at the Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, Italy. He held visiting research positions at Kyoto University, Japan, the University of Alicante, Spain, and the VTT Technical Research Center, Finland. Currently, he is an Assistant Professor with the Department of Mathematics, University of Padova, Italy. His research interests include machine learning and computer vision, with a particular focus on optimal transport and implicit neural learning, and applications to remote sensing data. He was awarded the MSCA-IF Fellowship for the OPTIMAL Project, which focuses on 3D point cloud change detection for looting analysis, from 2022 to 2024.

ARIANNA TRAVIGLIA (Associate Member, IEEE) received the Ph.D. degree in geomatics/ancient topography from the University of Trieste, Italy, in 2005. She held positions as a Postdoctoral Fellow and a Researcher at the University of Sydney and Macquarie University, Sydney, Australia, from 2006 to 2015. She lectured on "computing applications to archaeology and cultural heritage" and "computational thinking" at the University Ca' Foscari of Venice, from 2003 to 2018. Since 2019, she has been the Coordinator of the Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, Italy. Her research aims at developing digital practices within the study and management of cultural heritage, with expertise in multi- and hyperspectral image processing (close and far range). She has been the Coordinator, the PI, and the Co-PI of several projects funded by European Commission, European Space Agency (ESA), and Italian Space Agency (ASI), centered around the use of digital technologies and remote sensing data for cultural heritage analysis and protection. She has been the Executive Steering Committee Member of the International Computer Applications and Quantitative Methods in Archaeology (CAA) Association, the Co-Editor of *Journal of Computer Applications in Archaeology* (JCAA), and chaired the 41st CAA International Conference. She was a recipient of the H2020 MSCA-IF Fellowship, from 2015 to 2018.