Applied Computing and Geosciences 31 (2026) 100363

Contents lists available at ScienceDirect

Applied Computing and Geosciences

journal homepage: www.sciencedirect.com/journal/applied-computing-and-geosciences

Automated localization of landslide features on historical geological maps using deep learning

Abolfazl Moslemipour [a], Anna Małka [a,b,*], Frieder Enzmann [a], Jan Philip Hofmann [c], Michael Kersten [a]

[a] Institute of Geosciences, Johannes Gutenberg-University Mainz, Johann-Joachim-Becher-Weg 21, Mainz, D-55128, Germany

[b] Branch of Marine Geology, Polish Geological Institute–National Research Institute, 5 Kościerska Street, Gdańsk, 80-328, Poland

[c] State Office for Geology and Mining Rhineland-Palatinate, Emy-Roeder-Straße 5, Mainz, D-55129, Germany

ARTICLE INFO

Keywords:
Artificial intelligence
Deep learning
Oriented object detection
Dual-scale symbol localization
4-Channel input
Geological maps

ABSTRACT

The existence of small dimensions, thin lines, background noise, and colour changes due to paper wear makes manual symbol localization difficult on old geological maps. A method that addresses longstanding challenges in archival map utilization, such as collating heterogeneous symbols in old cartographic materials, is essential. Therefore, artificial intelligence can be beneficial in this field. By automating symbol localization, it unlocks previously underutilized archives for modern hazard studies, particularly in areas with limited contemporary monitoring. This study uses deep learning models and introduces a novel technique to automate the oriented localization of landslide-related symbols across a geological map series with a scale of 1:25,000 from the archive of the State Office for Geology and Mining Rhineland-Palatinate, Germany. Initially, geological images were cropped into tiles of two sizes: 128 × 128 and 416 × 416 pixels. Then, using established deep learning models, automatic localization of symbols using oriented bounding boxes (OBB) was performed at a single-scale and dual-scale by combining the detections at two scales. The YOLOv8 and v11 OBB models exhibited the best accuracy and detection performance among the deep learning models. Moreover, it was found that the dual-scale detection method achieved higher overall accuracy compared to the single-scale models, although it required significantly higher runtime. Furthermore, the 416 × 416-pixel tile size provided both lower runtime and higher detection accuracy than the 128 × 128-pixel tile, indicating that larger tiles enable more complete object representation and more efficient inference. Finally, a new 4-channel input method was introduced and used in the YOLOv11 OBB architecture. A detailed analysis of 94 historical and contemporary maps spanning 129 years (1887–2016) using this novel technique was able to provide accuracy at least at the same level as the dual-scale model. Therefore, the 4-channel model is considered the most optimal and efficient configuration for map symbol localization.

1. Introduction

Serial geological maps demonstrate the cultural and scientific significance of archival records. Since the late 19th century, German state authorities have systematically produced detailed geological maps, often with a scale of 1:25,000, supporting scientific research, resource exploration and mining, and state administration (Lippstreu, 2000). These archives document both geological knowledge and landscape changes, providing insights into areas now transformed by human activity. One of the most important and valuable applications of such archival data is the extraction and analysis of landslides. Anthropogenic landscape changes often hinder the clear identification of ancient landslide sites based solely on geomorphological features. Moreover, dormant landslides that occurred once in history tend to recur, underscoring the importance of archival information on mass movements, which is why such data are collected in many national and global landslide databases (Costa and Schuster, 1988; Grabowski et al., 2008; Korup, 2002). Information on past mass movements is crucial for determining the contemporary probability of landslides and for the creation of landslide hazard and risk maps (Corominas and Moya, 2008). However, the high relevance of archival data for landslide risk management is not reflected in the actual use of such historical data (Małka

* Corresponding author. Institute of Geosciences, Johannes Gutenberg-University Mainz, Johann-Joachim-Becher-Weg 21, D-55128 Mainz, Germany.
E-mail address: amalka@uni-mainz.de (A. Małka).

https://doi.org/10.1016/j.acags.2026.100363
Received 18 November 2025; Received in revised form 21 May 2026; Accepted 4 June 2026
Available online 6 June 2026

et al., 2016). This is primarily due to the sheer volume of collected data, which exceeds the analytical capacity of individual researchers. The compilation and localization of archival data are labour-intensive and do not always guarantee the retrieval of essential or valuable information.

Artificial intelligence (AI) has enabled the automated localization of geospatial features from historical maps, making these data more accessible and useable for a wide range of applications (Chen et al., 2019; Li et al., 2020; Sun et al., 2022). In an increasingly digital world, AI is playing a transformative role across multiple domains (Ding et al., 2024; Li et al., 2025; Moslemipour et al., 2025a, 2025b; Sadeghnejad et al., 2026; Xie et al., 2025). Deep learning networks can process large volumes of heterogeneous geospatial data, identifying patterns and relationships that may be overlooked by manual methods (Zhu et al., 2017). The use of deep learning for landslide detection in archival imagery has demonstrated high accuracy and transferability across different geographic regions, further supporting its utility for global and regional studies (Ghorbanzadeh et al., 2019). Historical records often reveal spatial and temporal clustering of landslides, frequently associated with extreme rainfall or seismic events (Crozier, 2010; Kirschbaum et al., 2015). Moreover, long-term landslide inventories are essential for probabilistic hazard modelling and spatial susceptibility mapping (Guzzetti et al., 2012; Van Den Eeckhaut and Hervás, 2012). Combining archival data with contemporary satellite and LiDAR datasets allows for multi-temporal analysis of landscape evolution and the potential for landslide reactivation, providing valuable insights into the dynamics of geohazards (Malamud et al., 2004; Tanyaş et al., 2017). Therefore, Geospatial AI (GeoAI) integrates AI techniques with geospatial data to enable advanced analysis and prediction of geological phenomena across geographic space (Hu et al., 2024; Janowicz et al., 2020; Lunga et al., 2022; Mai et al., 2025; Reichenbach et al., 2018). GeoAI techniques enable automated localization of map features from diverse data, supporting both current and historical cartographic analysis at various scales (Mai et al., 2025). GeoAI has also rapidly become a transformative force in geological mapping, enabling unprecedented accuracy and efficiency in the analysis of large-scale remote sensing datasets (Hu et al., 2024; Janowicz et al., 2020; Lunga et al., 2022).

However, the symbolic representation of lithology and surface phenomena, as well as the historical development of these symbols since the 19th century, present unique challenges for AI techniques. Contemporary landslide localization research predominantly focuses on satellite and drone imagery (Bhuyan et al., 2023; Chen et al., 2014; Meena et al., 2023; Schoenfeldt et al., 2022; Zhang et al., 2022), while automated localization of historical landslide symbols on geological maps remains mostly unexplored. Modern deep learning networks such as convolutional neural networks (CNNs) and the You Only Look Once (YOLO) family of algorithms, widely recognized for their efficiency and accuracy in real-time applications (Redmon et al., 2016), enable efficient image detection and pattern identification in this context (Jocher and Qiu, 2024). In more recent work, Wang et al. (2025) introduced a deep learning model for landslide detection in remote sensing images. By adding the Receptive Field Attention (RFA) module, using the Normalized Wasserstein Distance (NWD), and simplifying the YOLOv11 architecture, they demonstrated that the accuracy of detecting small and multi-scale landslides can be significantly improved. Guo et al. (2021) introduced a two-stage deep learning framework for geological symbol detection. First, individual symbols are identified with the single symbol detection network (SSDN), and then composite symbols are extracted with a novel graph convolutional network with L2 distance attention (DAGCN). Moreover, Saxton et al. (2024) presented a novel method for automatically extracting polygonal and point features from historical maps, utilizing open-set segmentation and detection and using map units as prompts. However, previous studies have mainly focused on modern topographic symbols using high-quality contemporary data. These studies generally do not consider the difficult conditions and characteristics of historical maps, including poor quality scans, ink stains, paper discoloration, folding, tearing, water stains, print displacement, and moisture warping (Huang et al., 2023; Zhang et al., 2022). Therefore, introducing faster and more accurate methodologies for symbol localization in such maps is essential.

In this study, the potential of deep learning models is explored for localizing landslides and related features on historical and contemporary 1:25,000 geological maps (GM25) of the German state of Rhineland-Palatinate using oriented bounding boxes (OBB). This approach aims to overcome the highly labour-intensive nature of manual analysis and demonstrates how deep learning can localize landslide-related symbols even from old and degraded cartographic materials of the 19th and 20th centuries. For this purpose, the geological maps were initially split into three parts: training, validation, and testing. Then, the images were cropped into smaller tiles in two different sizes (128 and 416 pixels). Then, using oriented deep learning networks including RetinaNet OBB, Faster RCNN OBB, Faster RCNN ORPN, Oriented Former, YOLOv8 OBB and YOLO11 OBB, training was carried out separately on each dataset. Finally, single- and dual-scale detections were performed to analyse the accuracy of geological symbol detection. To further improve the accuracy of the deep learning model while reducing computational cost, a novel 4-channel input configuration was introduced. This configuration incorporated the three standard RGB channels along with an additional distance transform channel derived from multi-scale edge information. The inclusion of this fourth channel enhanced both the precision and recall of the model and, most importantly, significantly reduced the runtime compared to the dual-scale approach.

2. Materials and methods

2.1. Map history and selection

The implementation of deep learning in this study is based on the examination of detailed historical maps of Rhineland-Palatinate geology at a scale of 1:25,000 as the primary data source. Geological maps at this scale (GM25) are the most effective means of representing rock formations on the Earth's surface, utilizing symbols and colours for geological features and phenomena visualization. The production of large-scale official maps has been a core responsibility of national geological surveys (Małka et al., 2016). Understanding the characteristics and historical context of detailed geological maps in Germany is essential for addressing the challenges of deep learning-based analysis. All original analogue GM25 maps are currently preserved in the archives of the State Office for Geology and Mining Rhineland-Palatinate (LGB). This study focuses on maps published between 1887 and 2016. The oldest studied geological maps were the Schillingen (6306) and Hermeskeil (6307) GM25 sheets, prepared and published in 1887. In total, 94 of the geological maps found in this repository were analysed. The analysed maps encompass regions adjacent to large urban centres, highlighting areas particularly vulnerable to landslides and their potentially severe impact on urban populations and infrastructure (Fig. 1).

Most of these maps are accompanied by supplementary textual explanations. The extensive time span (>129 years), during which these GM25 sheets were produced, led to multiple changes in their format, content, layout, and symbology, influenced by evolving economic needs and advances in the geosciences (e.g., changes in stratigraphic classification and nomenclature). The GM25 sheets were compiled during different mapping periods by various authors, each reflecting the lithological, lithogenic, stratigraphic, and nomenclatural standards of their respective era. Additionally, the affiliation of the responsible institution underwent several changes, further affecting the map content (Fig. 1). The first map series published for this German state region was the ‘Geological Special Map of Prussia and the Thuringian States’, compiled and issued by the Royal Prussian Geological Survey. For the present-day territory of the Rhineland-Palatinate state, 37 sheets from this series

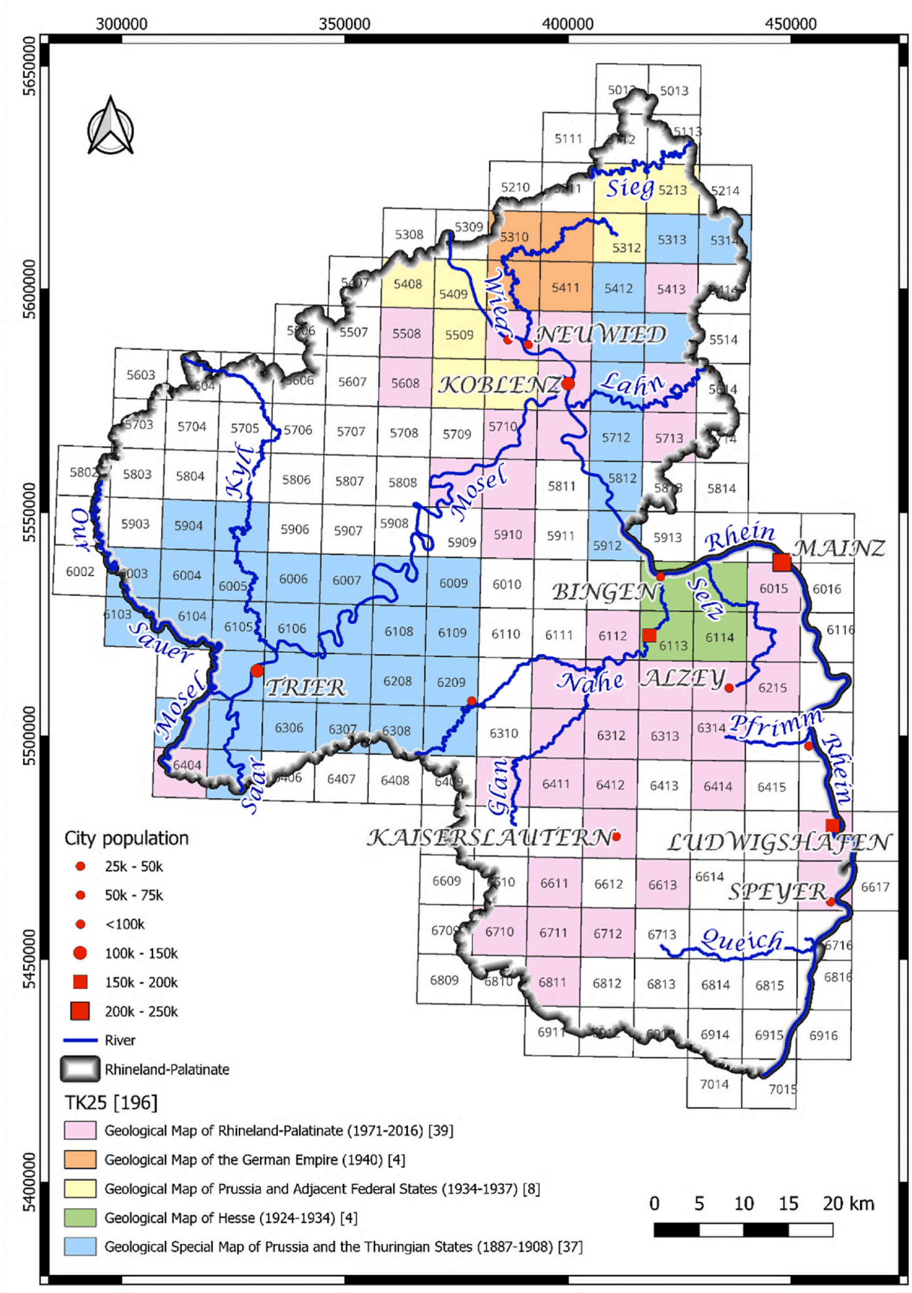


Fig. 1. Age structure of 1:25,000 geological map sheets of Rhineland-Palatinate used for deep learning.

were published over a span of 21 years (1887–1908). Subsequently, geological map production was interrupted for an extended period due to World War I and ensuing political changes. The next relevant series, the ‘Geological Map of Prussia and Adjacent Federal States’, was issued by the Prussian Geological Survey, with 8 sheets published between 1934 and 1937. Additionally, between 1925 and 1934, four sheets were published under the title ‘Geological Map of Hesse’. Between 1939 and 1945, geological maps were released under the designation ‘Geological Map of the German Empire’, including four sheets published in 1940, although these were originally surveyed between 1928 and 1931. Geological cartography during this era was significantly hindered by the geopolitical disruptions of both World Wars. Following World War II, there was another prolonged hiatus in state-sponsored geological mapping. The production of new geological maps was resumed in the 1970s, under the title ‘Geological Map of Rhineland-Palatinate’, with a total of 39 sheets published between 1971 and 2016.

Clearly, the evolution of geological mapping in the region was strongly influenced by both historical and political developments, as well as the resulting administrative changes. Since state authorities financed the production of geological maps, the extent of mapping at a scale of 1:25,000 varied considerably over time, leading to uneven coverage and persistent gaps. The survey of the long and complex history of map production reveals a heterogeneous and varied cartographic corpus of significant territorial scope. This heterogeneity extends to the base topographic information, thematic scope, content quality, geoscientific markings, colour schemes, and the diverse geological symbols and conventions employed over different periods. Despite these challenges, it is evident that historical Prussian geological and topographic maps represent a valuable source of information on the highly diverse geomorphological and geological structures of the Rhineland-Palatinate region, reflected in various types of mass movements.

2.2. Dataset preparation

Due to the long and complex history of GM25 production, different conventions and symbols were used for the same phenomena across various periods. For deep learning-based analysis, features relevant to landslide research were prioritized, and symbol selection was guided by the mass movement literature (Dikau, 1996; Hungr et al., 2014; Varnes, 1984). A template index of key symbol characteristics from different periods was then created for training. Fig. 2 illustrates the relevance of geological symbols utilized in deep learning.

In this study, an annotation toolbox on the online platform CVAT (Computer Vision Annotation Tool; https://www.cvat.ai/) was used to annotate GM25 images and generate a high-quality dataset for training and evaluating deep learning models in geological symbol detection. All GM25 images were manually annotated with oriented bounding boxes around geological symbols and assigned corresponding class labels. In total, 12,164 individual symbols were labelled into 12 classes as shown in Fig. 3. Some symbols, for example ‘landslide’, ‘spring’ and ‘mine’ (Fig. 4), have more than one shape on the geological maps and, therefore, a separate class and type were defined for each shape.

All 94 geological map images were divided into three dataset parts, as common with deep learning tools: training, validation, and testing, with a ratio of 85%, 10%, and 5%, respectively. Therefore, 80 map images were used for training, 9 map images were allocated for validation, and 5 map images were considered for the test dataset. One important challenge is that the geological maps are much larger than the symbols on them; the maps are typically about 6100 × 5700 pixels in size, whereas most symbols are smaller than 100 × 100 pixels. To address these challenges, the tiling technique was used. Table 1 shows

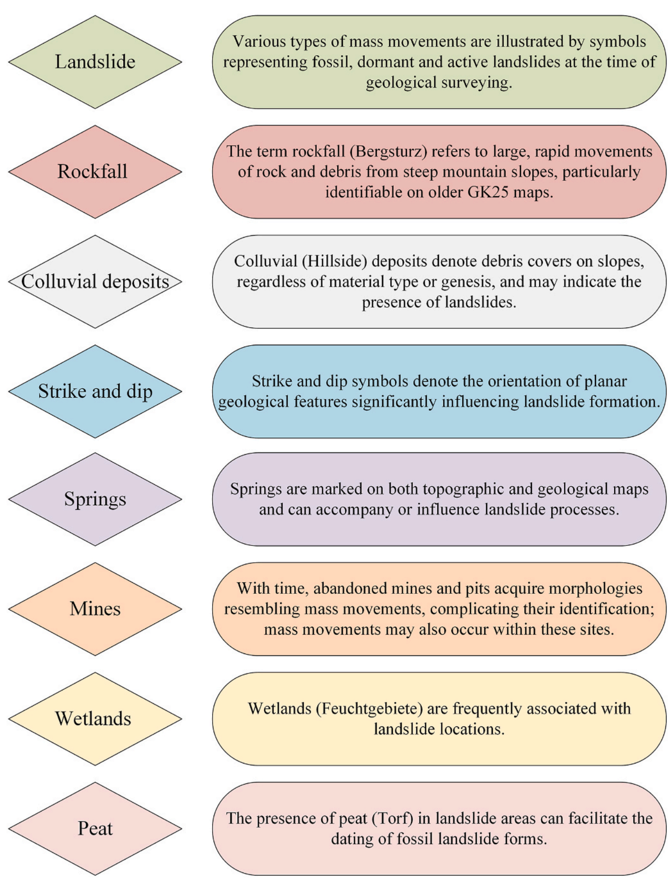


Fig. 2. Geological symbols used for analysing a 1:25,000 scale geological map (GM25) series for the Rhineland-Palatinate state of Germany.

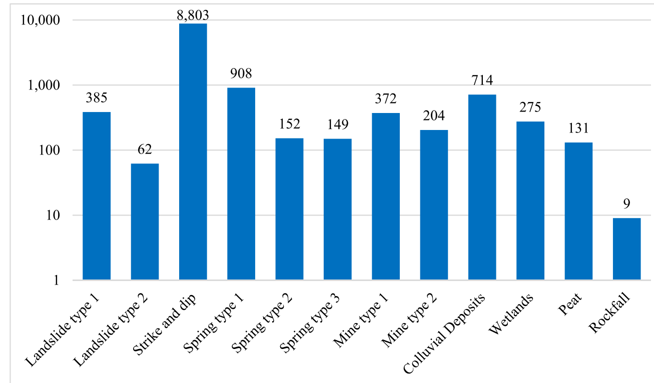


Fig. 3. Class-number statistics for the ground truth after annotation.

the sizes of the annotated labels in the ground truth images. Clearly, the minimum length and width of the labels are about 7 pixels, and their maximum length and width are approximately 344 and 211 pixels, respectively. On the other hand, about 95% of the labels have a length smaller than 80.9 pixels and a width smaller than 46 pixels (P95). Therefore, considering the size of the labels, a size of 128 pixels was selected for image training and symbol detection. However, since only a few symbols have a length or width greater than 128 pixels, a size of 416 (which is larger than the maximum label dimension of 344 pixels) was also used during training and detection to ensure that all symbols could be effectively captured and compared.

Therefore, each geological map was cropped into tiles of two sizes: 416 × 416 pixels and 128 × 128 pixels, with overlaps of 100 pixels and 30 pixels, respectively. The cropping workflow is illustrated in Fig. 5. Fig. 5a represents an original geological map, which is subdivided into smaller tiles in Fig. 5b. Fig. 5c highlights the overlapping region, ensuring that symbols located near tile boundaries are also included in the training data. Following the cropping phase, a series of data augmentation techniques were applied to increase the number of objects for symbol classes with fewer than 800 annotations (Shorten and Khoshgoftaar, 2019). Random scaling was applied to the map tiles by a fixed factor of 1.2. Additionally, random horizontal and vertical translations of ±30 pixels were introduced, with corresponding updates to label positions. Maps were also converted to the HSV colour space and modified by randomly adjusting brightness and saturation levels, ranging between 0.6 and 1.4.

For 128 × 128-pixel tiles with a 30-pixel overlap, a total of 252,753 image tiles were generated for the training data, of which 14,692 tiles had labels and 238,061 were empty. After data augmentation, the number of labelled tiles increased to 26,282, and 105,136 of the empty tiles were kept for training (1:4 ratio). For the validation data, a total of 31,411 image tiles were generated, of which 1547 tiles had labels and 29,864 had no labels, of which 6188 were kept (1:4 ratio). The test images were similarly divided into 14,198 small tiles, of which 530 had labels and 13,668 had no labels. Moreover, for 416 × 416-pixel tiles with a 100-pixel overlap, a total of 22,607 image tiles were created for the training data, of which 7308 tiles had labels and 15,299 were empty. After data augmentation, the number of labelled tiles increased to 15,900. For the validation data, a total of 2839 image tiles were created, of which 911 tiles had labels and 1928 had no labels. The test images were also divided into 1292 small tiles, of which 369 had labels and 923 had no labels.

2.3. Object detection models

In this research, after tiling and augmenting the images to prepare the dataset, six different object detection models were employed. These include three single-stage detectors, two two-stage detectors, and one End-to-End Transformer-based detector, specifically designed for

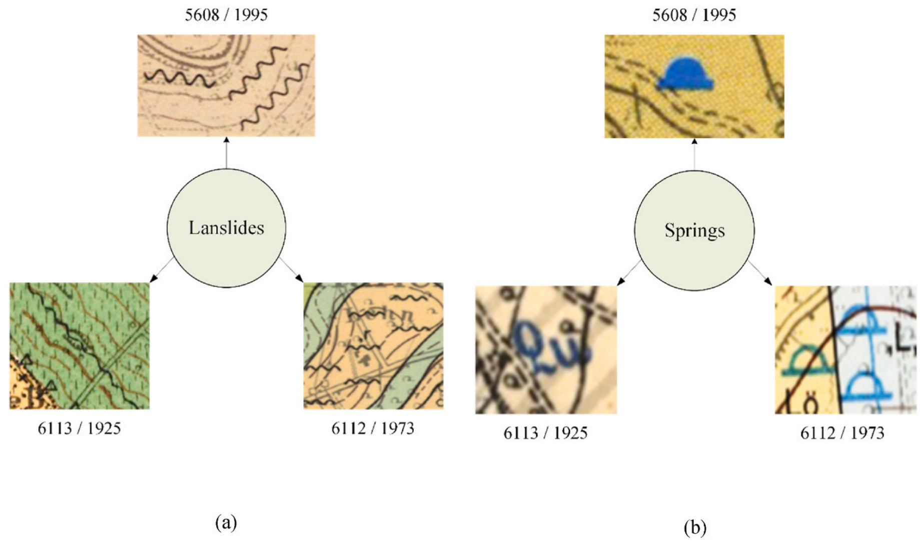


Fig. 4. Examples of the different symbol shapes for (a) landslides and (b) springs from the GM25. The numbers below each crop indicate the sheet number and name/year of GM25 publication.

Table 1
Symbol size overview in annotated geological maps.

	Minimum	Mean	Maximum	P95
Length (pixels)	7.60	46.06	344.4	80.9
Height (pixels)	6.34	24.37	211.1	46.0

oriented (directional) detection of symbols in geological images. The models are YOLOv11 OBB (Jocher and Qiu, 2024), YOLOv8 OBB (Jocher et al., 2023), Faster RCNN OBB (Ren et al., 2015), RetinaNet OBB (Lin et al., 2017), Faster RCNN ORPN (Xie et al., 2024), and Oriented Former (Zhao et al., 2024). Each of these models is described in detail in the Supporting Material. The implementations of Faster RCNN OBB, RetinaNet OBB, and Faster RCNN ORPN were based on the source code provided in https://github.com/jbwang1997/OBBDetection.

2.4. 4-Channel input

The main challenge in automatic localization of geological symbols is their small dimensions, thin lines, background noise, and colour changes due to paper wear. These factors prevent standard object detection models that rely on RGB input from properly separating symbol features from noise and background. To overcome this limitation, a 4-channel input mechanism was designed and implemented in this study. In this method, in addition to the three main image channels (RGB), one additional channel was generated and added to the YOLOv11 model. The fourth channel was automatically generated based on a combination of gradient information and edge-distance information (Distance Transform of Multi-Scale Edges). For this purpose, first, the input image was filtered at several different scales with sigma values (0.0, 0.8, 1.6, and 3.2) to create a multi-scale gradient map. Then, percentage thresholding with values ($p_{hi} = 90$) and ($p_{lo} = 65$) was used to extract edge areas, and then morphological opening operations were applied to remove local noise. Next, the distance transform was calculated on non-edge areas, and normalization was performed using the smooth function ($\exp(-d/\tau)$) with a value of $\tau = 3.0$. Finally, by combining 70% of the smoothed distance map and 30% of the initial gradient intensity, a continuous and smooth edge map (DT-Edge) was obtained, which was added as a fourth channel to the three main RGB channels. The computational parameters of DT-Edge were determined empirically based on qualitative visual evaluations of a subset of validation images containing stains, tears, paper discolouration, and very small symbols. Sigma values (0.0, 0.8, 1.6, 3.2) were selected to extract boundaries of varying thicknesses: smaller values are necessary to identify fine, thin symbol lines, whereas larger values capture thicker edges while suppressing fine local noise. The percentage threshold ($p_{hi} = 90$, $p_{lo} = 65$) values were set at this level to ensure that only the strongest gradients associated with the original ink of the symbols were retained, while weaker gradients caused by old paper texture, creases, and background stains were effectively filtered out. Moreover, the smoothing parameter value ($\tau = 3.0$) was selected to provide a sufficient penetration radius to connect gaps and faded lines in old maps, ensuring that symbols remain unified without causing the boundaries of adjacent symbols to blend or blur. Thus, each input image had a structure of (R, G, B, DT-Edge) with dimensions (H, W, 4), and this combination increased the sensitivity of the YOLO model to the boundaries and textures of objects in complex data. Moreover, an alternative configuration was designed in which the edges extracted with the Canny algorithm were added to the YOLOv11 model as an additional channel alongside the three RGB channels. This 4-channel Canny-based input (R, G, B, Canny) was implemented solely as a comparison configuration alongside the 4-channel DT-Edge-based input.

2.5. Model training and detection

In this study, deep learning models were trained on two datasets independently. In YOLOv8 and v11, for the 128 × 128-pixel tile dataset, training parameters included a batch size of 64 over 150 epochs, using the SGD optimizer with a learning rate of 0.005 and weight decay of 0.001. For the 416 × 416-pixel tile dataset, a batch size of 16 over 150 epochs was also employed, but with a learning rate of 0.002 and weight decay of 0.003. Moreover, early stopping was enabled for both models. Box loss, classification loss, and distribution focal loss were used to evaluate the error in bounding box detections. Box loss quantifies the error in the predicted bounding box coordinates, thereby measuring the model's accuracy in localizing objects. Classification loss evaluates the accuracy with which the model assigns correct class labels to detected objects. Additionally, distribution Focal Loss was applied to further refine bounding box predictions and enhance localization precision (Li et al., 2020). All training, evaluation, and inference phases were conducted using the PyTorch v2.6.0 deep learning framework (Paszke et al., 2019).

For other deep learning models, training was also conducted using tile sizes of 128 × 128 and 416 × 416 pixels separately. First, both

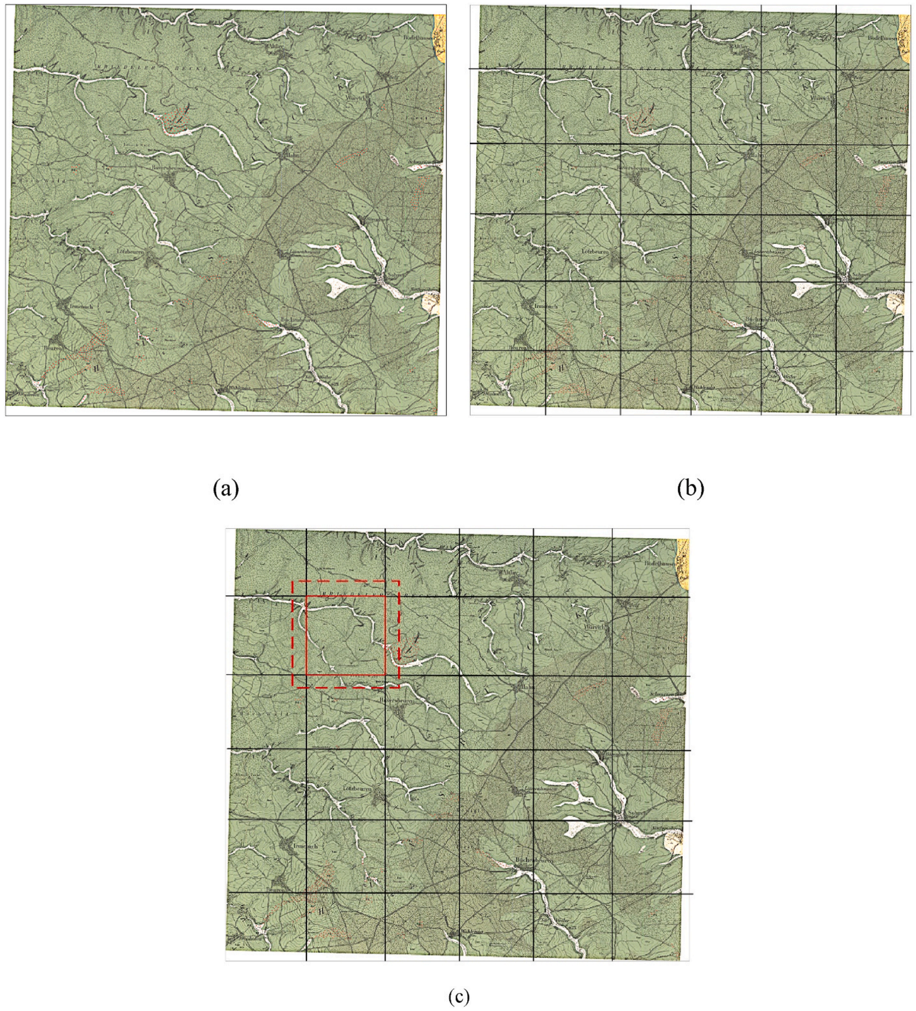


Fig. 5. Cropping process of geological map images with a specified overlap. (a) The original map image. (b) Different tiles after cropping with a specified tile size. (c) The overlap area around tile.

datasets were converted to DOTA (Dataset for Object Detection in Aerial Images). Subsequently, RetinaNet was trained separately using the SGD optimizer, with an initial learning rate of 0.003, momentum of 0.9, and a weight decay of 0.0001. A stepwise learning rate schedule was adopted, with a linear warmup (ratio = 0.001) applied during the first 500 iterations. The learning rate decreased at epochs 24 and 33, and training continued for a total of 36 epochs. The batch size was set to 8 for the 416 $\times$ 416-pixel dataset and 16 for the 128 $\times$ 128-pixel dataset. Likewise, for the two-stage detector networks (Faster RCNN and Faster RCNN ORPN) and the End-to-End Transformer-based detector (Oriented Former), training followed the same procedure with the SGD optimizer and an initial learning rate of 0.003. Moreover, momentum (0.9), weight decay (0.0001), learning rate warmup, stepwise policy, and decay periods (epochs 24 and 33) remained the same. Training also lasted 36 epochs, with batch sizes of 8 for 416 $\times$ 416 tiles and 16 for 128 $\times$ 128 tiles.

In this study, the oriented object detection process was performed using all deep learning models developed for multi-class detection based on tiled detection. In the single-scale method, the input image is divided into overlapping tiles, and each tile is fed separately to the trained deep learning model (e.g., 128 or 416 pixels). The output of the model consists of oriented bounding boxes that are merged at the global scale after removing duplicate areas and applying a confidence threshold of 0.25. In the dual-scale method, detection is performed simultaneously at two different tile sizes. Then the results are merged using a cross-scale consensus filter, so that only the detections having sufficient overlap and confidence in both scales are retained. Finally, the output is saved by drawing oriented boxes on the image and recording the coordinates, angle, and confidence value of each object. All the training and detection were performed on a workstation equipped with a 32-core Intel® Xeon® Platinum 8160 CPU (2.1 GHz), 1 TB RAM, and dual Nvidia GeForce RTX 2080 Ti GPUs (11 GB VRAM each). Furthermore, to comprehensively assess the performance of the object detection models, a range of evaluation metrics were employed in this study, which are described in the Supporting Material.

3. Results and discussion

3.1. Symbol detection

Symbol localization was first performed separately on each tile size and then by merging the detections performed on each tile, the dual-scale detection results were also reported. Fig. 6a, b, and 6c show the results for the 128 $\times$ 128, 416 $\times$ 416, and dual tile sizes, respectively. Moreover, Table S1 (the Supporting Material) summarizes the complete precision, recall, F1-score, and mAP results for all detections, along with their discussion. These results indicate that in geological applications, where the goal is simply to detect the presence of a feature rather than its precise delineation, YOLO models, especially YOLOv8 in the dual-scale configuration, offer the best balance between accuracy and coverage by being able to understand the shape and angle of the features without depending on anchors. However, not only accurate but also fast symbol detection is a critical requirement in automated geological map analysis,

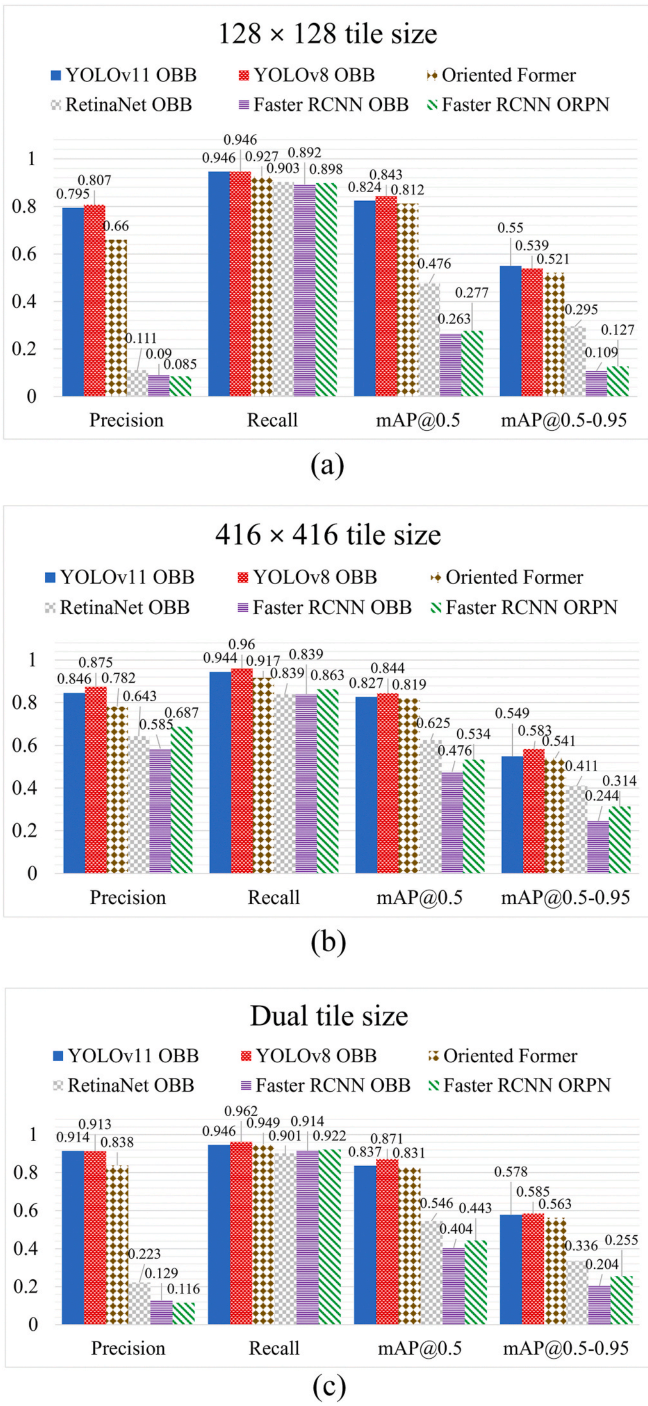


Fig. 6. Test performance comparison between different deep learning models in (a) 128 × 128, (b) 416 × 416, and (c) dual tile size symbol detections.

and recent advances in deep learning architectures have significantly improved both detection speed and precision (Redmon et al., 2016).

Table 2 shows the runtime for detecting all symbols in an image with a size of 6160 × 5730 pixels. From this table, we can see that the YOLOv8 and YOLOv11 models had the lowest runtime for detecting all symbols among all the models examined. These two models, with a single-stage detector and lighter architecture, perform on average about 4 to 5 times faster than anchor-based models and about 3 times faster than Transformer-based models. At a tile size of 128 × 128, YOLOv8 identified all symbols on a map sheet in only 37.3 s, and YOLOv11 in 50.4 s, while RetinaNet required 144.2 s, Faster RCNN required 179.7 s, and Oriented Former required 124.6 s for the same processing. Even the more optimized version of ORPN, at 169.3 s, is still more than three times slower than YOLOv8. At the more efficient 416 × 416 tile size, the runtime is significantly reduced, as fewer tiles are processed to cover the entire map. At this tile size, YOLOv11 was the fastest model at 10.2 s and YOLOv8 at 10.4 s. In contrast, RetinaNet, Faster RCNN, ORPN, and Oriented Former took 21.8 s, 21.6 s, 21.5 s, and 17.2 s, respectively. In the dual-scale tile size, although the runtime increases (as shown in the Supporting Material sections on the quantitative test results), the accuracy of the model also increases significantly. Running the two 128 and 416 tile models in combination takes 61.2 s and 47.2 s for YOLOv11 and YOLOv8, respectively. However, this time is still much less than that of RetinaNet (167.4 s), Faster RCNN (202.7 s), ORPN (191.6 s), and Oriented Former (139.8 s). In practical comparison, if a sheet of geological map is manually reviewed by an expert, the time required is usually a full day. In contrast, YOLO models can analyse the same sheet with acceptable accuracy in only about 10 to 60 s. Therefore, these models are not only superior in terms of detection accuracy but also in terms of time efficiency, enabling the fast and automatic extraction of large-scale map symbols.

Table 2
Symbol detection runtime of different deep learning models for various tile sizes.

Deep learning model	Run time (Seconds)		
	128 × 128 tile size	416 × 416 tile size	Dual tile size
YOLOv11 OBB	50.4	10.2	61.2
YOLOv8 OBB	37.3	10.4	47.2
Oriented Former	124.6	17.2	139.8
RetinaNet OBB	144.2	21.8	167.4
Faster RCNN OBB	179.7	21.6	202.7
Faster RCNN ORPN	169.3	21.5	191.6

3.2. YOLOv11 OBB 4-channel input

For training and testing 4-channel input data, the 416 × 416 tile size of the YOLOv11 OBB model was selected. Considering its metric values and lower runtime, this model is a suitable option for further improving symbol localization accuracy on geological maps. As shown in the results, the dual-scale YOLOv8 and v11 models have the highest metric values for symbol detection, and using the information at two scales allows several FPs to be detected and removed from the final detection. However, both YOLOv8 and v11 models with the 416 × 416 tile size showed the lowest runtime, requiring only about 10 s for an image in a size of 6160 × 5730 pixels to detect symbols, while using dual-scale detection for the same image requires about 50 to 60 s and is therefore 5 to 6 times slower. Additionally, using dual-scale detection requires training on two tiles of different sizes, which also increases the computational cost and runtime.

Fig. 7 shows the convergence and stability of the YOLOv11 OBB models in the 3-channel and 4-channel versions. In both cases, the box loss (blue) and classification loss (red) decrease rapidly in the first few epochs and then approach a stable value with a gentle slope, indicating proper convergence of the network and the absence of overfitting. In the 4-channel version (Fig. 7a and d), the initial values of the losses are slightly lower, and the decrease is more uniform, especially in the val_cls_loss section, where the fluctuations after epoch 20 are smaller than in the 3-channel case, indicating that the addition of a fourth channel has led to greater stability and better context understanding in the network. Besides, the dfl_loss, representing the accuracy of box localization is almost constant in both models, but with a slight decreasing trend, and remains at a value of around one.

In addition, Fig. 8 shows the relative percentage changes in the performance indicators of the YOLOv11 OBB model for the 4-channel and dual-scale modes compared to the base 3-channel model. In Fig. 8a, which corresponds to the validation data, the 4-channel model

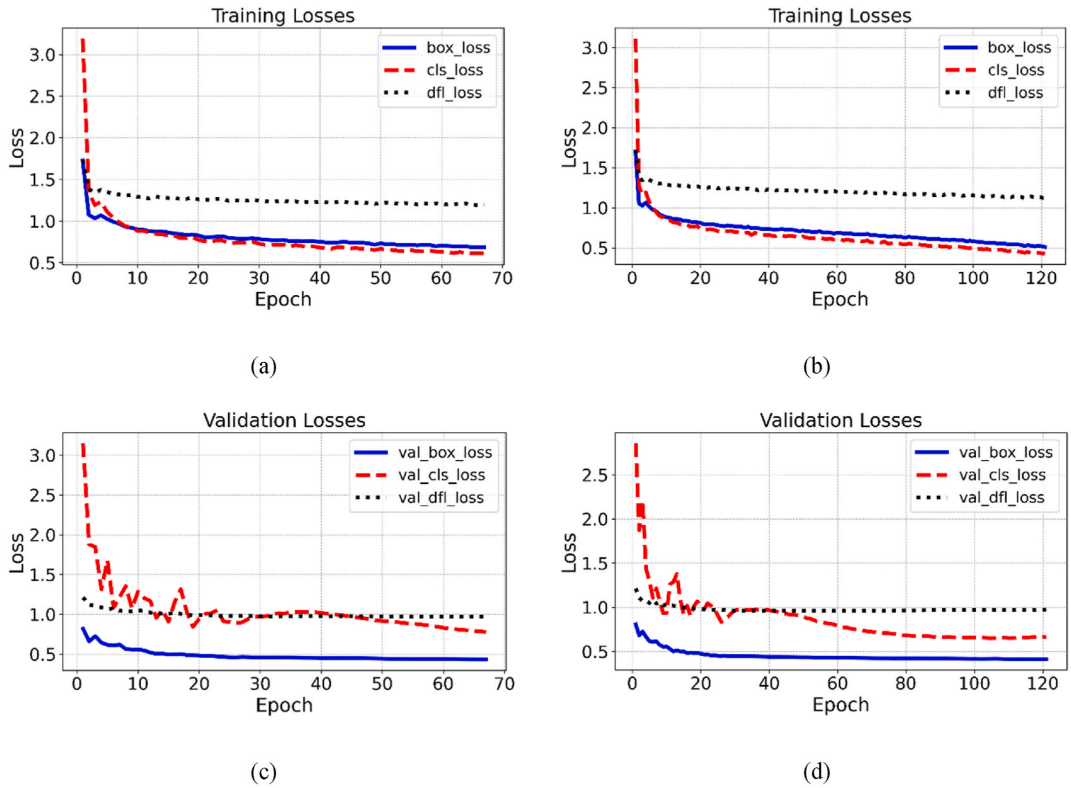
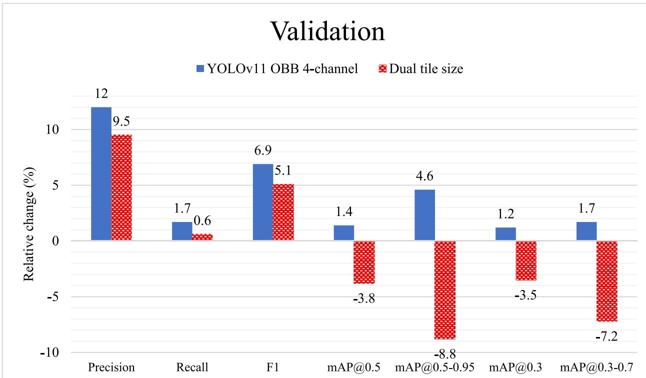


Fig. 7. Training loss curves for (a) YOLOv11 OBB 3-channel and (b) YOLOv11 OBB 4-channel, along with validation loss curves for (c) YOLOv11 OBB 3-channel and (d) YOLOv11 OBB 4-channel.

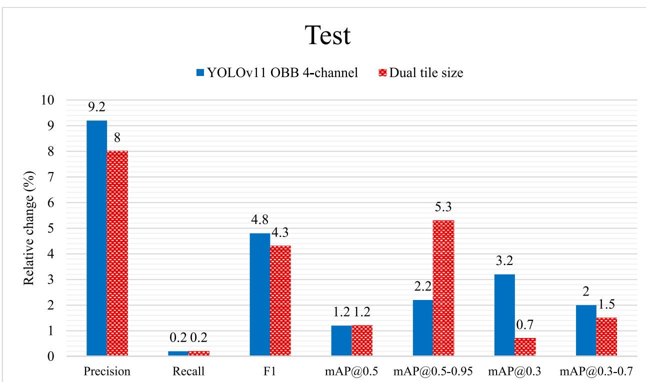
shows improvements across all metrics compared to the 3-channel model. Its precision, recall, and F1-score values have increased by about 12%, 1.7%, and 6.9%, respectively. The mAP metrics have also grown in all IoU intervals, including mAP@0.5 by 1.4% and mAP@0.5–0.95 b y 4.6%, compared to the 3-channel model. Even in the softer indices mAP@0.3 and mAP@0.3–0.7, there is an increase of 1.2% and 1.7%, respectively, which indicates that the model is able to detect symbols more accurately by adding a fourth channel of information. From a scientific point of view, this improvement is due to the addition of textural and spectral information in the fourth channel, which allows the network to gain a better understanding of the local context and texture of the image. Therefore, it achieves more accurate performance in distinguishing between subtle symbols and the same-coloured background, and the model can detect objects of different sizes without the need for multiple scales. Moreover, the dual-scale model performs better than the 3-channel model in some areas, especially in precision (9.5% increase), recall (0.6% increase), and F1-score (5.1% increase). However, unlike the 4-channel model, there is a decrease in mAP metrics. The mAP@0.5 value has decreased by 3.8% and mAP@0.5–0.95 b y 8.8%, and the softer indices such as mAP@0.3 (−3.5%) and mAP@0.3–0.7 (−7.2%) have also decreased. These decreases, despite the overall increase in accuracy, occurred due to the mismatch between the predictions at the two scales (128 and 416). From a technical point of view, when small objects are detected at both scales, boxes with high overlap but different spatial alignment are produced. In the fusion and NMS application stage, this mismatch can lead to over-suppression or partial integration of some valid predictions, mainly affecting the spatial accuracy at high IoUs and consequently reducing the mAP, while the recall value does not change much (only +0.6%).

In Fig. 8b, corresponding to the test data, there is a subtle but important difference in the behaviour of the two models. In this dataset, the dual-scale model, unlike the validation dataset, shows a significant improvement in the more accurate mAP indices; mAP@0.5–0.95 increases by 5.3% and mAP@0.3–0.7 by 1.5%, while precision, recall, and F1 score are also about 8.0%, 0.2%, and 4.3% higher than those of the 3-channel model, respectively. The scientific reason for this difference between validation and test datasets is seen in the statistical composition of the data. In the test data, the symbols are usually located separately from each other and with a more balanced aspect ratio, and as a result, merging the results of the two scales does not cause double elimination. Therefore, mAP increases in the test dataset, but in the validation dataset, which has many closely spaced objects and boundaries, this effect is reversed. Meanwhile, the 4-channel model in the test data has a uniform and stable improvement in all metrics and outperforms both the 3-channel and dual-scale models in almost all indicators. This model shows a 9.2% increase in precision, 0.2% increase in recall, 4.8% in F1-score, and about 2-3% in all mAPs (mAP@0.5 = +1.2%, mAP@0.5–0.95 = +2.2%, mAP@0.3 = +3.2%, mAP@0.3–0.7 = +2%). The reason for this stability is that the addition of the fourth channel enriches the feature space and the network can learn a wide range of symbol sizes and orientations in a single forward pass without depending on multiple scales. As a result, unlike the dual-scale model, boundary errors are eliminated, and the model performs more stably in both soft and hard IoUs.

Table 3 compares the TP, FP, and FN values for the YOLOv11 OBB model with RGB, RGB + DT-Edge, RGB + Canny, and dual tile inputs in the validation and test datasets. The results show that adding Canny-based binary edges to the RGB input not only did not improve performance, but also increased detection errors, as the number of FPs increased from 195 to 216 in the validation set and from 64 to 72 in the test dataset. This Canny-based 4-channel input also decreased the number of TPs from 960 to 953 in the validation set and from 351 to 347 in the test dataset. Canny edges are a binary, local representation of intensity gradients that lack any contextual or structural information about the continuity and width of boundaries. In historical maps, heterogeneous colour variations, stains, cracks, and paper folding create gradients similar to real boundaries; as a result, the Canny channel produces a large number of spurious edge responses and reinforces



(a)



(b)

Fig. 8. Relative percentage changes of detection performance metrics for the YOLOv11 OBB 4-channel and dual tile size configurations with respect to the YOLOv11 OBB 3-channel baseline on the validation (a) and test (b) datasets.

Table 3
Detection performance comparison (TP, FP, and FN) across different YOLOv11 OBB configurations on the validation and test datasets.

Dataset	Count	YOLOv11 OBB			
		RGB	RGB + DT-Edge	RGB + Canny	Dual tile
Validation	**TP**	960	976	953	966
	FP	195	72	216	95
	FN	101	85	108	95
Test	**TP**	351	352	347	352
	FP	64	29	72	33
	FN	21	20	25	20

irrelevant features in the early layers of the network. In contrast, it clearly shows that the DT-Edge fourth channel significantly reduced FPs while maintaining a high detection rate. In the validation set, the 3-channel model has 960 TP, 195 FP, and 101 FN, while the 4-channel model has performed more consistently and accurately with 976 TP, only 72 FP, and 85 FN. In other words, adding the DT-Edge fourth channel has resulted in a more than 60% reduction in FP without any reduction in actual detection. The DT-Edge channel provides a continuous, area-based representation of the feature structure by encoding the distance of each pixel to the nearest extracted boundary at multiple scales. The use of multi-scale gradient information enhances edges that are stable at different scales and implicitly carries information about the thickness, continuity, and relative importance of boundaries, while suppressing unstable responses due to local noise. In addition, applying distance transformation and exponential smoothing preserves the spatial continuity of incomplete or blurred boundaries, allowing them to be interpreted as coherent spatial patterns. The dual-scale model also performs close to the DT-Edge-based 4-channel model with 966 TP, 95 FP, and 95 FN, indicating a slight increase in the loss of detection of small or borderline objects. A similar trend was observed in the test data in that RGB, RGB + DT-Edge, and dual tile inputs have almost the same number of TPs (351 to 352), but the differences in FP are significant; the 3-channel model has recorded 64 FP, the DT-Edge 4-channel only 29 FP, and the dual-scale 33 FP. At the same time, the FN value in the DT-Edge 4-channel and dual-scale models is equal and minimal (20), which indicates the maximum detection rate. Therefore, from an analytical point of view, the DT-Edge 4-channel model has established a more optimal balance between sensitivity and accuracy than the 3-channel and even dual-scale models by significantly reducing FP and maintaining TP.

Furthermore, the runtime difference is significant. The 3-channel model takes about 10.2 s, the DT-Edge 4-channel model takes 13.9 s, and the dual-scale model takes 61.2 s to detect symbols in a full image. Therefore, while the dual-scale model improves accuracy at the cost of more than a six-fold increase in runtime, the DT-Edge 4-channel model performs similarly or better than the dual-scale model in most metrics, with only 3.7 s more runtime than the single-scale 3-channel model. This suggests that the DT-Edge 4-channel YOLOv11 is a more reliable and cost-effective alternative to the dual-scale model, especially in practical geological mapping applications that require fast, stable, and scalable analysis.

3.3. Visual representation

As mentioned, the DT-Edge 4-channel and dual-scale detection models have been shown to significantly enhance object detection performance, particularly in complex visual environments with variable object sizes and backgrounds (Chen et al., 2019). Fig. 9 presents selected detection results on some images, illustrating the effectiveness of the deep learning models. The detection outputs demonstrate both the diversity and accuracy of symbol identification across different map styles, colour schemes, and textures. All symbols are detected with high confidence scores, frequently exceeding 0.85 and reaching up to 1.00 for certain symbols.

Small symbols like 'springs' and 'mines', as well as elongated or context-dependent features such as 'landslide' and 'strike and dip', are consistently detected with precise boundaries. The use of multiple colours and label classes in the output further indicates that the model reliably distinguishes between different symbol categories, maintaining high category integrity. Additionally, the model-oriented symbol detection capability allows it to estimate the strike angle for features such as 'strike and dip' as illustrated in Fig. 9e. Moreover, analysis of historical geological maps enables the identification and determination of the timing of formation or existence (at least since) for numerous landslides not recorded in the landslide database (LDB). For example, the landslide northeast of Nieder-Olm (Fig. 10), absent from the LDB, is clearly visible on a 3D microtopographic map and has existed since at least the map's publication in 1989.

3.4. Discussion

In this study, the choice of OBBs is a conscious geometric simplification and is not intended to replace precise polygonal demarcations of landslides. In many conventional approaches to landslide susceptibility analysis, precise polygons are used as input, while OBB-based representation inevitably includes non-landslide areas. Therefore, the direct use of this type of data has limitations for precise quantitative analyses without additional processing. However, in the context of historical geological maps, the precise boundaries of landslides are often not clearly discernible, i.e. the highly general and poorly delineated content is usually represented symbolically, particularly in older maps. In such situations, the extraction of precise polygons can lead to vague interpretations and depend on expert judgment. Consequently, it may convey a degree of certainty that does not exist in the underlying data. In

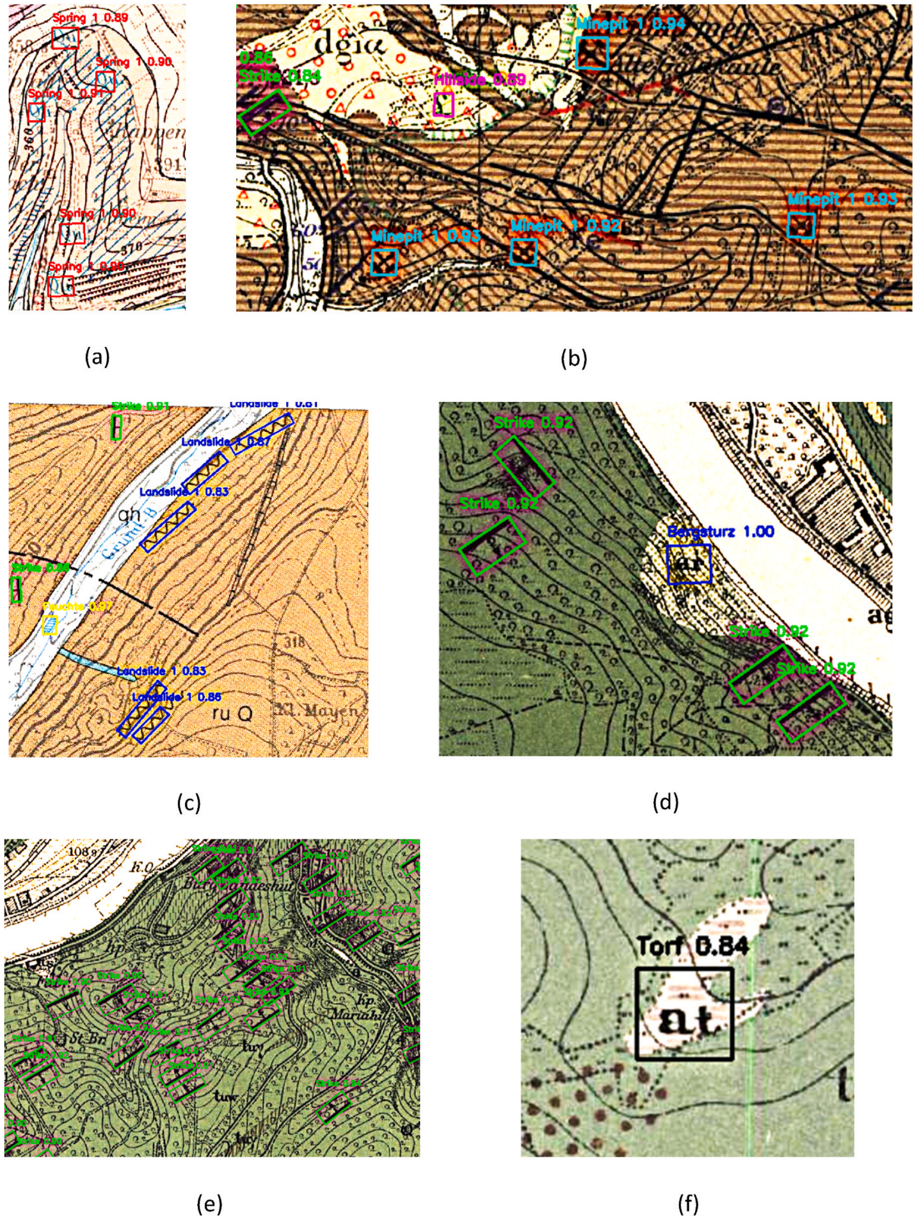


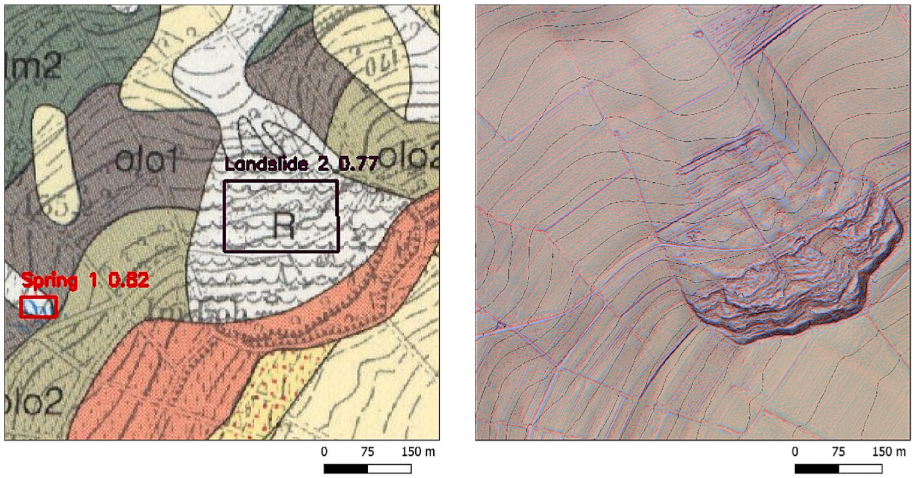
Fig. 9. Examples of predicted symbols on some geological maps. Each example primarily represents: (a) spring type 1, (b) mine type 1 and hillside deposits, (c) strike and dip, landslide type 1 and 'Feuchtgebiete' (wetland), (d) 'Bergsturz' (rockslide), strike and dip, (e) strike and dip, (f) 'Torf' (peat) symbols, respectively.

contrast, the OBB-based representation provides a stable and minimal description of the approximate location, orientation, and extent of the phenomenon, which is more compatible with the symbolic nature of historical maps. Accordingly, OBB-based data can be considered as an initial input layer for updating landslide databases, initial screening of susceptible areas, and detection of landslides not recorded in existing sources. These data can be converted into more precise polygonal boundaries in subsequent stages, in combination with modern data (such as LiDAR or satellite imagery) or expert review.

Due to the nature of the historical maps used, the absolute spatial accuracy of the extracted results is affected by several factors. These maps have usually undergone physical changes over time, such as folding, shrinkage, or local deformation, and in addition, the scanning and georeferencing process can lead to heterogeneous spatial distortions on the map surface. In such circumstances, common machine vision metrics, such as mAP, although suitable for evaluating the performance of detection and relative localization in image space, do not indicate the true spatial accuracy in geographic coordinates. Accordingly, the results of this study should be interpreted in the context of relative and consistent localization in image space, and not as a precise metric representation of the location of features on the ground. Especially in areas close to the edges of the map, there is a possibility of increased spatial errors due to distortions, which can affect the final location of the extracted features.

4. Conclusions

This study demonstrates the transformative potential of deep learning models in analysing historical geological maps for landslide research. By applying different deep learning networks including YOLOv11 OBB, YOLOv8 OBB, RetinaNet OBB, Faster RCNN OBB, Faster RCNN ORPN, and Oriented Former to 94 GM25 maps (1:25,000 scale) from Rhineland-Palatinate published between 1887 and 2016, we achieved the following results.



(a) (b)

Fig. 10. Example of a landslide northeast of Nieder-Olm: (a) on the geological map, sheet 6015 Mainz, (b) the same landslide visualized using a 3D microtopographic map computed according to the method of (Miyagi et al., 2024).

- The analysis of 94 geological map sheets demonstrated that deep learning reduced processing time from a full day (manual) to around 10 s per sheet, facilitating rapid handling of large archival collections. This acceleration supports comprehensive studies of landslide occurrence and distribution.
- Due to their anchor-free structure, the YOLOv8 and v11 models performed better than Transformer-based models as well as classical models including RetinaNet, Faster RCNN and Faster RCNN ORPN in detecting small and oriented symbols on geological maps.
- Increasing the tile size from 128 $\times$ 128 to 416 $\times$ 416 resulted in a steady improvement in precision (about +8%) and an increase in the average mAP@0.5–0.95 by +5%, as objects are seen more completely at larger scales and the models gain better spatial accuracy.
- The use of a dual tile size increased the accuracy on the test data (specifically mAP@0.5–0.95 $\approx$ +5.3%), but on the validation data, mAP dropped (about −8%). This decrease is due to the lack of perfect alignment of the predictions in the fusion step from the two different scales. Therefore, the performance of this method depends on the statistical composition of the data and object density.
- The DT-Edge 4-channel YOLOv11 model outperformed the 3-channel version in almost all metrics and provided similar or better performance than the dual-scale model without the need for two independently trained networks.
- From a computational perspective, the DT-Edge 4-channel model's runtime (13.9 s) was only slightly longer than the single-scale model (10.2 s) and more than four times faster than the dual-scale model (61.2 s). Therefore, this model has the best balance between accuracy, stability, and computational cost.
- Based on TP, FP, and FN counts, the DT-Edge 4-channel model had more than 60% FP reduction compared to the 3-channel model in the validation data (72 vs. 195), while the TP value even increased (976 vs. 960), indicating a sharp reduction in FPs without any loss in TP.
- Overall, the DT-Edge 4-channel YOLOv11 model is considered the most suitable option for automatic detection of geological map symbols in terms of overall precision, recall, stability across different datasets, and runtime, and can be an efficient and low-cost alternative to dual-scale methods or heavy anchor-based networks.

In future studies, the proposed detection method can be combined with georeferencing processes to allow for quantitative assessment of absolute spatial accuracy. In particular, alignment of historical maps with modern reference data such as satellite imagery or digital elevation models can help reduce heterogeneous distortions caused by scanning and physical deformation of maps. Moreover, despite the relative robustness of the DT-Edge channel to noise and degradation, which are common in historical maps, a quantitative and systematic assessment of the impact of different types of physical degradation, including paper folding, discolouration, stains, and scanning distortions, could be conducted in future studies. Analysing the behaviour of the fourth channel under different levels of degradation and comparing it with binary edge-based methods could lead to a more accurate understanding of the stability of this representation and its further optimization for highly degraded archival data. Moreover, the visual stability analysis showed that the algorithm is robust for the GM25 archive analysed in this study. However, applying it to other historical maps with different digitisation quality may require a slight fine-tuning of these parameters.

Statement

During the preparation of this work, the authors used Perplexity Pro to assist with literature review and language editing. After using Perplexity Pro, the authors thoroughly reviewed and edited all content as necessary and take full responsibility for the accuracy and integrity of the publication.

Computer code availability

Name of the code: Oriented Object Detection.

Contact: amoslemi@uni-mainz.de.

Hardware requirements: Minimum: standard PC or laptop with 8 GB RAM and optional CUDA-compatible GPU.

Tested on: a workstation equipped with 32 Intel® Xeon® Platinum 8160 2.1 GHz CPU cores, 1 TB of RAM, and two 11 GB Nvidia GeForce RTX 2080 Ti GPUs.

Program language: Python 3.12.

Software required: ultralytics==8.3.196, opencv-python==4.11.0.86, numpy==1.26.4, shapely==2.0.7, openpyxl==3.1.5, pandas==2.3.1.

Program size: ~15 MB (including source code, sample input/output images, and documentation).

The source codes are available for downloading at the following link: https://github.com/Abolfazlmsl/Oriented-Object-Detection.

CRediT authorship contribution statement

Abolfazl Moslemipour: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Anna Małka:** Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Frieder Enzmann:** Conceptualization, Funding acquisition, Resources, Supervision. **Jan Philip Hofmann:** Formal analysis, Writing – review & editing. **Michael Kersten:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study has been funded by the MABEIS III project ('Mass Movement Information System Rhineland-Palatinate'), funded by the State Office for Mobility Rhineland-Palatinate (LBM) and the State Office for Geology and Mining Rhineland-Palatinate (LGB). The authors would like to acknowledge Felix Hettwer for his assistance in preparing the city data for Fig. 1. The authors would also like to thank Adam Trubłajewicz and Anita Gerhard for help with proofreading

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.acags.2026.100363.

Data availability

All dataset that supports the findings of this study are available from the corresponding author upon reasonable request.

References

Bhuyan, K., Tanyaş, H., Nava, L., Puliero, S., Meena, S.R., Floris, M., Van Westen, C., Catani, F., 2023. Generating multi-temporal landslide inventories through a general deep transfer learning strategy using HR EO data. Sci. Rep. 13, 162. https://doi.org/10.1038/s41598-022-27352-y.

Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., 2019. MMDetection: open mmlab detection toolbox and benchmark. arXiv preprint arXiv:1906.07155. https://doi.org/10.48550/arXiv.1906.07155.

Chen, Y., Lin, Z., Zhao, X., Wang, G., Gu, Y., 2014. Deep learning-based classification of hyperspectral data. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 7, 2094–2107. https://doi.org/10.1109/JSTARS.2014.2329330.

Corominas, J., Moya, J., 2008. A review of assessing landslide frequency for hazard zoning purposes. Eng. Geol. 102, 193–213. https://doi.org/10.1016/j.enggeo.2008.03.018.

Costa, J.E., Schuster, R.L., 1988. The formation and failure of natural dams. Geol. Soc. Am. Bull. 100, 1054–1068. https://doi.org/10.1191/0309133302pp333ra.

Crozier, M.J., 2010. Deciphering the effect of climate change on landslide activity: a review. Geomorphology 124, 260–267. https://doi.org/10.1016/j.geomorph.2010.04.009.

Dikau, R., 1996. Landslide Recognition: Identification, Movement and Courses (No Title).

Ding, L., Chen, B., Zhu, Y., Dong, H., Chan, G., Zhang, P., 2024. Geo-Hgan: unsupervised anomaly detection in geochemical data via latent space learning. Comput. Geosci. 192, 105703. https://doi.org/10.1016/j.cageo.2024.105703.

Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S.R., Tiede, D., Aryal, J., 2019. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. Remote Sens. 11, 196. https://doi.org/10.3390/rs11020196.

Grabowski, D., Marciniak, P., Mrozek, T., Nescieruk, P., Rączkowski, W., Wójcik, A., 2008. Instrukcja Opracowania Mapy Osuwisk i Terenów Zagrożonych Ruchami Masowymi W Skali 1:10 000. Ministerstwo Środowiska/Państwowy Instytut Geologiczny, Warszawa.

Guo, M., Bei, W., Huang, Y., Chen, Z., Zhao, X., 2021. Deep learning framework for geological symbol detection on geological maps. Comput. Geosci. 157, 104943. https://doi.org/10.1016/j.cageo.2021.104943.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T., 2012. Landslide inventory maps: new tools for an old problem. Earth Sci. Rev. 112, 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001.

Hu, Y., Goodchild, M., Zhu, A.-X., Yuan, M., Aydin, O., Bhaduri, B., Gao, S., Li, W., Lunga, D., Newsam, S., 2024. A five-year milestone: reflections on advances and limitations in GeoAI research. Ann. Gis 30, 1–14. https://doi.org/10.1080/19475683.2024.2309866.

Huang, W., Sun, Q., Yu, A., Guo, W., Xu, Q., Wen, B., Xu, L., 2023. Leveraging deep convolutional neural network for point symbol recognition in scanned topographic maps. ISPRS Int. J. GeoInf. 12, 128. https://doi.org/10.3390/ijgi12030128.

Hungr, O., Leroueil, S., Picarelli, L., 2014. The Varnes classification of landslide types, an update. Landslides 11, 167–194. https://doi.org/10.1007/s10346-013-0436-y.

Janowicz, K., Gao, S., McKenzie, G., Hu, Y., Bhaduri, B., 2020. Geoai: Spatially Explicit Artificial Intelligence Techniques for Geographic Knowledge Discovery and Beyond. Taylor & Francis, pp. 625–636.

Jocher, G., Chaurasia, A., Qiu, J., 2023. Ultralytics YOLOv8.

Jocher, G., Qiu, J., 2024. Ultralytics YOLO11.

Kirschbaum, D., Stanley, T., Zhou, Y., 2015. Spatial and temporal analysis of a global landslide catalog. Geomorphology 249, 4–15. https://doi.org/10.1016/j.geomorph.2015.03.016.

Korup, O., 2002. Recent research on landslide dams-a literature review with special attention to New Zealand. Prog. Phys. Geogr. 26, 206–235. https://doi.org/10.1191/0309133302pp333ra.

Li, Q., Wang, Z., Li, W., Hu, J., Rong, X., Bian, L., Wu, W., 2025. Object segmentation of near surface magnetic field data based on deep convolutional neural networks. Comput. Geosci. 196, 105847. https://doi.org/10.1016/j.cageo.2024.105847.

Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., Tang, J., Yang, J., 2020. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection. Adv. Neural Inf. Process. Syst. 33, 21002–21012.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal loss for dense object detection. Proceedings of the IEEE International Conference on Computer Vision, pp. 2980–2988.

Lippstreu, L., 2000. Von den Anfängen der geologischen Kartierung im ehemaligen Preußen–ein Beitrag zum Beginn der geologisch-agronomischen Kartierung im Norddeutschen Flachland vor 125 Jahren. Brand. Geowiss. Beitrage 7, 5–19.

Lunga, D., Hu, Y., Newsam, S., Gao, S., Martins, B., Yang, L., Deng, X., 2022. GeoAI at ACM SIGSPATIAL: the new frontier of geospatial artificial intelligence research. SIGSpatial Special 13, 21–32. https://doi.org/10.1145/3578484.3578491.

Mai, G., Xie, Y., Jia, X., Lao, N., Rao, J., Zhu, Q., Liu, Z., Chiang, Y.-Y., Jiao, J., 2025. Towards the next generation of Geospatial Artificial Intelligence. Int. J. Appl. Earth Obs. Geoinf. 136, 104368. https://doi.org/10.1016/j.jag.2025.104368.

Małka, A., Jegliński, W., Relisko-Rybak, J., 2016. Prussian geological maps of Northern Poland in the archives of the Polish Geological Institute and their current application in geology. Pol. Cartogr. Rev. 48. https://doi.org/10.1515/pcr-2016-0017.

Malamud, B.D., Turcotte, D.L., Guzzetti, F., Reichenbach, P., 2004. Landslide inventories and their statistical properties. Earth Surf. Process. Landf. 29, 687–711. https://doi.org/10.1002/esp.1064.

Meena, S.R., Nava, L., Bhuyan, K., Puliero, S., Soares, L.P., Dias, H.C., Floris, M., Catani, F., 2023. HR-GLDD: a globally distributed dataset using generalized deep learning (DL) for rapid landslide mapping on high-resolution (HR) satellite imagery. Earth Syst. Sci. Data 15, 3283–3298. https://doi.org/10.5194/essd-15-3283-2023.

Miyagi, T., Ikeda, K., Ishikawa, H., Doan, L., Thanh, N.K., Van Tien, P., Li, Y., Zhang, F., 2024. Interpretation and Mapping for the Prediction of Sites at Risk of Landslide Disasters: from Aerial Photography to Detection by Dtms, Progress in Landslide Research and Technology, vol. 3. Springer, pp. 15–61. https://doi.org/10.1007/978-3-031-55120-8_2. Issue 1, 2024.

Moslemipour, A., Sadeghnejad, S., Enzmann, F., Khoozan, D., Hupfer, S., Schäfer, T., Kersten, M., 2025a. Multi-scale pore network fusion and upscaling of microporosity using artificial neural network. Mar. Petrol. Geol. 177, 107349. https://doi.org/10.1016/j.marpetgeo.2025.107349.

Moslemipour, A., Sadeghnejad, S., Enzmann, F., Khoozan, D., Schäfer, T., Kersten, M., 2025b. Efficient multi-scale image reconstruction of heterogeneous rocks with unresolved porosity using octree structures. Comput. Geosci. 29, 24. https://doi.org/10.1007/s10596-025-10362-w.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., 2019. Pytorch: an imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. 32.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788.

Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., Guzzetti, F., 2018. A review of statistically-based landslide susceptibility models. Earth Sci. Rev. 180, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001.

Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: towards real-time object detection with region proposal networks. Adv. Neural Inf. Process. Syst. 28. https://doi.org/10.1109/TPAMI.2016.2577031.

Sadeghnejad, S., Hasanloo, A., Moslemipour, A., Schäfer, T., 2026. Automated identification of stylolites in geological whole-core images using hybrid deep learning networks. Applied Computing and Geosciences, 100335. https://doi.org/10.1016/j.acags.2026.100335.

Saxton, A., Dong, J., Bode, A., Jaroenchai, N., Kooper, R., Zhu, X., Kwark, D.H., Kramer, W., Kindratenko, V., Luo, S., 2024. Accurate feature extraction from historical geologic maps using open-set segmentation and detection. Geosciences 14, 305. https://doi.org/10.3390/geosciences14110305.

Schoenfeldt, E., Winocur, D., Pánek, T., Korup, O., 2022. Deep learning reveals one of Earth's largest landslide terrain in Patagonia. Earth Planet Sci. Lett. 593, 117642. https://doi.org/10.1016/j.epsl.2022.117642.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. Journal of big data 6, 1–48. https://doi.org/10.1186/s40537-019-0197-0.

Sun, Z., Sandoval, L., Crystal-Ornelas, R., Mousavi, S.M., Wang, J., Lin, C., Cristea, N., Tong, D., Carande, W.H., Ma, X., 2022. A review of earth artificial intelligence. Comput. Geosci. 159, 105034. https://doi.org/10.1016/j.cageo.2022.105034.

Tanyaş, H., van Westen, C.J., Allstadt, K.E., Anna Nowicki Jessee, M., Görüm, T., Jibson, R.W., Godt, J.W., Sato, H.P., Schmitt, R.G., Marc, O., 2017. Presentation and analysis of a worldwide database of earthquake-induced landslide inventories. J. Geophys. Res. Earth Surf. 122, 1991–2015. https://doi.org/10.1002/2017JF004236.

Van Den Eeckhaut, M., Hervás, J., 2012. State of the art of national landslide databases in Europe and their potential for assessing landslide susceptibility, hazard and risk. Geomorphology 139, 545–558. https://doi.org/10.1016/j.geomorph.2011.12.006.

Varnes, D.J., 1984. Landslide Hazard Zonation: a Review of Principles and Practice.

Wang, B., Su, J., Xi, J., Chen, Y., Cheng, H., Li, H., Chen, C., Shang, H., Yang, Y., 2025. Landslide detection with MSTA-YOLO in remote sensing images. Remote Sens. 17, 2795. https://doi.org/10.3390/rs17162795.

Xie, X., Cheng, G., Wang, J., Li, K., Yao, X., Han, J., 2024. Oriented R-CNN and beyond. Int. J. Comput. Vis. 132, 2420–2442. https://doi.org/10.1007/s11263-024-01989-w.

Xie, Z., Li, K., Jiang, J., Yang, J., Qiao, X., Yuan, D., Nie, C., 2025. An integrated neighborhood and scale information network for open-pit mine change detection in high-resolution remote sensing images. Comput. Geosci. 196, 105880. https://doi.org/10.1016/j.cageo.2025.105880.

Zhang, H., Zhou, X., Li, H., Zhu, G., Li, H., 2022. Machine recognition of map point symbols based on YOLOv3 and automatic configuration associated with POI. ISPRS Int. J. GeoInf. 11, 540. https://doi.org/10.3390/ijgi11110540.

Zhao, J., Ding, Z., Zhou, Y., Zhu, H., Du, W.-L., Yao, R., El Saddik, A., 2024. OrientedFormer: an end-to-end transformer-based oriented object detector in remote sensing images. IEEE Trans. Geosci. Rem. Sens. https://doi.org/10.1109/TGRS.2024.3456240.

Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci. Remote Sens. Mag. 5, 8–36. https://doi.org/10.1109/MGRS.2017.2762307.