Journal of Applied Hydrography

HYDROGRAPHISCHE NACHRICHTEN 06/2026 HN 134

DOI: 10.23784/HN134-02

AI-driven detection of quartzite blocks in the River Rhine using CNN U-Net with backscatter and bathymetric data

An article by ERIC IDUN

Submerged quartzite blocks in shallow inland waterways present a significant hazard to navigation, particularly in high-traffic rivers such as the River Rhine. Conventional detection methods rely on manual interpretation of multibeam echo sounder (MBES) data, which is time-consuming, subjective and difficult to scale. This study evaluates the application of deep learning for automated detection of submerged rock features using bathymetric and backscatter data. A U-Net convolutional neural network with a ResNet-50 encoder was trained on a dataset comprising over 8,000 manually digitised rock features from a 7.5 km section of the Rhine. In addition to single bathymetry and backscatter inputs, derivative features were generated and combined using a principal component analysis to form composite datasets. Model performance was assessed using Intersection over Union (IoU), precision, recall and F1-score. The best results were achieved using a composite dataset integrating backscatter, slope and roughness, yielding an IoU of 0.492 and an F1-score of 0.659. The model demonstrated strong agreement with manual annotations and identified additional potential features not previously mapped. The results highlight the value of combining acoustic intensity and terrain derivatives in deep-learning workflows. This study demonstrates a scalable and objective approach to submerged hazard detection, with implications for improving hydrographic survey efficiency and navigational safety.

boulder detection | inland water mapping | autonomous data processing | backscatter data | CNN | PCA
Quarzitblockerkennung | Binnengewässerkartierung | autonome Datenverarbeitung | Rückstreudaten | CNN | PCA

Untergetauchte Quarzitblöcke in flachen Binnengewässern stellen eine erhebliche Gefahr für die Schifffahrt dar, insbesondere in stark befahrenen Flüssen wie dem Rhein. Herkömmliche Detektionsmethoden basieren auf der manuellen Auswertung von Fächerecholotdaten (MBES), was zeitaufwendig, subjektiv und schwer skalierbar ist. Diese Studie evaluiert die Anwendung von Deep Learning zur automatisierten Detektion untergetauchter Felsformationen mit Hilfe bathymetrischer Daten und Rückstreudaten. Ein U-Net-Faltungsnetzwerk mit einem ResNet-50-Encoder wurde mit einem Datensatz trainiert, der über 8000 manuell digitalisierte Felsformationen eines 7,5 km langen Rheinabschnitts umfasst. Zusätzlich zu den Rohdaten aus Bathymetrie und Rückstreuung wurden abgeleitete Merkmale generiert und mittels Hauptkomponentenanalyse zu zusammengesetzten Datensätzen kombiniert. Die Modellleistung wurde anhand von Intersection over Union (IoU), Präzision, Trefferquote und F1-Score bewertet. Die besten Ergebnisse wurden mit einem kombinierten Datensatz erzielt, der Rückstreuung, Neigung und Rauheit integrierte und einen IoU-Wert von 0,492 sowie einen F1-Score von 0,659 ergab. Das Modell zeigte eine hohe Übereinstimmung mit manuellen Annotationen und identifizierte zusätzliche, bisher nicht kartierte potenzielle Merkmale. Die Ergebnisse unterstreichen den Nutzen der Kombination von akustischer Intensität und Geländederivaten in Deep-Learning-Workflows. Diese Studie demonstriert einen skalierbaren und objektiven Ansatz zur Erkennung von Unterwassergefahren mit Implikationen für die Verbesserung der Effizienz hydrographischer Vermessungen und der Sicherheit der Schifffahrt.

Author
Eric Idun completed the Master programme Geodesy and Geoinformatics with specialisation in Hydrography at HafenCity University Hamburg.

ericidun92@gmail.com

Introduction

Safe navigation in inland waterways is critical to global trade, and rivers such as the Rhine play a central role in European transport networks. However, the dynamic nature of fluvial systems presents ongoing challenges for hydrographic surveying. Sediment transport, erosion and morphological changes continuously reshape the riverbed, leading to the exposure of submerged hazards. Quartzite blocks are among the most significant hazards in the Rhine (Boenigk 2021). Due to their high resistance to erosion, these rocks persist while surrounding sediments are removed, gradually becoming exposed. In shallow sections, they pose a direct risk to vessels, potentially causing hull damage or grounding.

Accurate and timely detection is therefore essential for maintaining navigational safety and updating nautical charts. Modern hydrographic surveys provide high-resolution bathymetric and backscatter data using MBES systems. However, interpretation of these datasets remains largely manual. This process is labour-intensive, time-consuming and subject to variability between operators. With increasing data volumes, manual workflows are becoming inefficient and difficult to sustain. Recent advances in artificial intelligence, particularly deep learning, offer new opportunities for automation. Convolutional neural networks (CNNs), and specifically U-Net architectures, have proven highly effective for image segmentation tasks (Cai et al. 2020; LeCun et al. 2015). Their ability to perform pixel-level classification makes them well suited for identifying complex patterns in geospatial data. This study investigates the use of a CNN-based approach to detect submerged quartzite blocks. The objectives are to evaluate model performance, assess the benefit of integrating bathymetric and backscatter features, and determine the most effective dataset configuration. The study contributes to hydrography by demonstrating a data-driven method for improving detection accuracy and efficiency in fluvial environments.

Data and study area

The study area (see Fig. 1) is situated in the lower Rhine region of Germany, specifically in Düsseldorf, and forms part of the W27 project area. This area extends between Rhine kilometres 722.5 (Düsseldorf) and 769.5 (Krefeld). It is divided into four legs (see Fig. 1) corresponding to prospective construction sites. The leg designated TA3, referred to as »Steinerne Bänke«, is located in the northern section of Düsseldorf and spans approximately 10 km (Kraft et al. 2025). This section was selected due to the notable concentration of rocks, making it particularly suitable for a detailed analysis. For the purposes of this study, the survey focused on the area between Rhine kilometres 747.0 and 754.5, located between the Theodor-Heuss-Brücke and the Langst-Kaiserswerth ferry.

Methodology

Data acquisition and processing

The dataset used in this study was acquired through hydrographic surveys conducted along a 7.5 km stretch of the River Rhine. Data acquisition was performed using a multibeam echo sounder (MBES) system capable of capturing high-resolution bathymetric measurements and co-registered backscatter intensity data. The MBES survey, conducted aboard the *DVocean* vessel, took place between 24 and 31 July 2024 during periods of mean water level to optimise acquisition conditions. Data acquisition was carried out using QPS Qinsy and Kongsberg SIS software. The multibeam echo sounder operated at a frequency of 300 kHz with a swath opening angle of 120°.

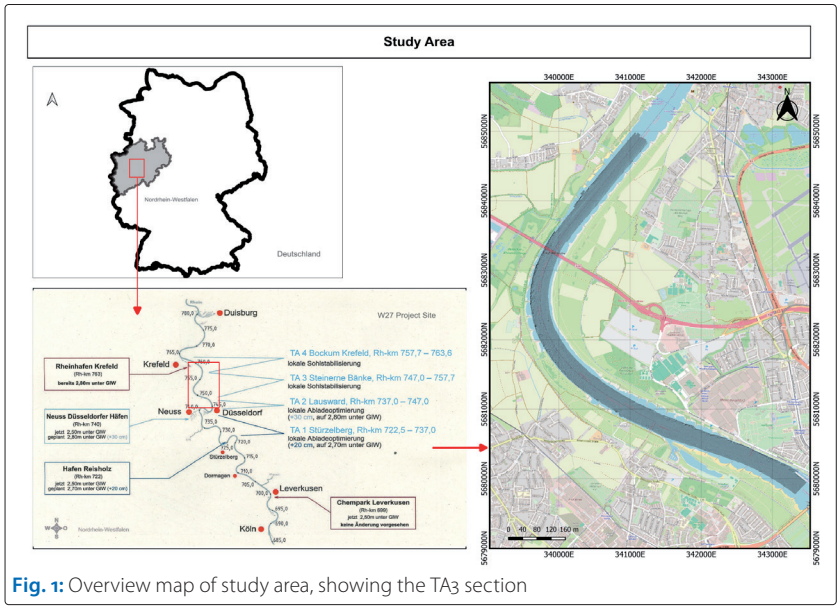

Fig. 1: Overview map of study area, showing the TA3 section

Bathymetric data was processed to remove noise, correct for motion, sound velocity and positioning errors, and gridded into DTMs. Backscatter data underwent radiometric corrections to account for transmission loss, beam pattern effects and angular dependency. In addition to the primary datasets, bathymetric derivatives such as slope and roughness were generated to capture terrain characteristics, while texture features such as contrast and dissimilarity were derived from backscatter data using Grey Level Co-occurrence Matrix (GLCM) methods. Features that visibly enhanced the detectability of rocky substrates were prioritised.

Manual digitising

Manual digitisation was conducted to create ground truth data for supervised learning. Rock features were identified through visual interpretation of both bathymetric and backscatter datasets. Characteristic indicators included elevated structures in bathymetry, high backscatter intensity, and associated morphological features such as scour marks. A total of over 8,000 rock features were digitised across the study area. These features were converted into binary masks, where each pixel was labelled as either rock or background. The resulting dataset provided a reference for model training and evaluation. While manual digitisation is essential for supervised learning, it introduces potential sources of uncertainty, including operator subjectivity and inconsistencies in feature delineation. These factors are considered in the interpretation of the model performance.

ID	Dataset	Training	Validation	Test	Total tiles	Channel
Ba	Bathymetry	279	163	23	465	Single band image
Bs	Backscatter	281	164	24	469	Single band image
BCD	Bathymetry + contrast + dissimilarity	253	148	21	422	Three band image
BsRS	Backscatter + roughness + slope	259	151	21	431	Three band image

Table 1: Number of tiles generated after pre-processing for each dataset

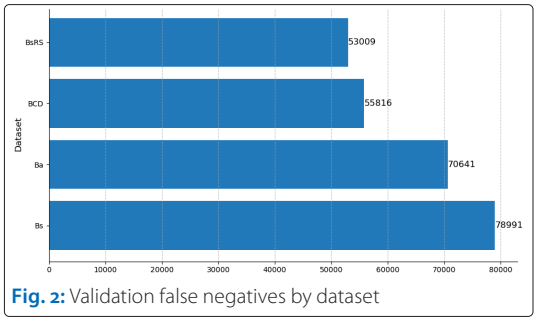

Fig. 2: Validation false negatives by dataset

Dataset pre-processing and preparation

Pre-processing was performed to prepare the dataset for model training. Principal component analysis (PCA) was applied as a feature selection approach to identify variables that best discriminate between rock formations. Rather than performing dimensionality reduction, PCA was used to evaluate feature importance based on component loadings. Variables with consistently high contributions to dominant components were retained. This process led to the construction of composite feature sets (see Table 1), including bathymetry–contrast–dissimilarity (BCD) and backscatter–roughness–slope (BsRS). Raster datasets were partitioned into tiles of 256 × 256 pixels to facilitate efficient training and capture local spatial patterns. All input features were normalised to ensure consistent scaling, and no-data regions were masked to avoid bias. Data augmentation techniques were applied to increase dataset diversity and improve generalisation. Class imbalance, due to the predominance of background pixels, was mitigated through sampling and training strategies.

Model training

The dataset comprised greyscale and multiband raster images with corresponding binary semantic segmentation masks. Model development was conducted using TensorFlow and the Keras API with the Segmentation Models library, and training was performed on Google Colab Pro with GPU acceleration. Models were trained using the Adam optimiser with an initial learning rate of 0.001 for up to 100 epochs and a batch size of 16. Binary cross-entropy loss was employed, given its suitability for pixel-wise binary classification tasks. Hyperparameters were selected based on empirical experimentation. Model performance was evaluated after each epoch using validation metrics.

Dataset	Validation Set Performance Metrics				Test Set Performance Metrics			
	IoU	Precision	Recall	F1 score	IoU	Precision	Recall	F1 score
Ba	0.3235	0.5174	0.4632	0.4888	0.3796	0.5470	0.5537	0.5503
Bs	0.2598	0.438	0.3897	0.4124	0.1141	0.2732	0.1639	0.2049
BCD	0.3841	0.5334	0.5785	0.5550	0.3875	0.5135	0.6122	0.5586
BsRS	0.4459	0.6044	0.6296	0.6167	0.4916	0.6148	0.7104	0.6592

Table 2: Comparative evaluation of model performance across datasets

Results

Across all datasets, models showed rapid initial convergence followed by stabilisation. However, generalisation performance varied significantly depending on input features. The BsRS dataset demonstrated the most stable training behaviour, with minimal overfitting and consistent validation loss.

The quantitative evaluation (see Table 2) shows that BsRS achieved the best overall performance (IoU: 0.4916, F1: 0.6592), followed by BCD and Ba. The Bs dataset performed poorly, particularly on the test set, indicating weak generalisation.

Error analysis

Confusion matrix analysis revealed that all models performed well in identifying background pixels, but detection of rock features varied. BsRS achieved the highest true positive rates and lowest false negatives (see Fig. 2), indicating strong sensitivity. In contrast, Bs exhibited high false negatives, suggesting difficulty in distinguishing rock features using backscatter alone.

Best performing dataset

The BsRS dataset (see Fig. 3) consistently outperformed all other inputs across training behaviour, quantitative metrics and qualitative assessment. Its superior performance is attributed to the integration of structural (slope, roughness) and radiometric (backscatter) features, enabling improved discrimination of rock features. In contrast, the Bs dataset showed limited performance due to ambiguity in backscatter signals, which can represent both rock and non-rock surfaces. Overall, BsRS was

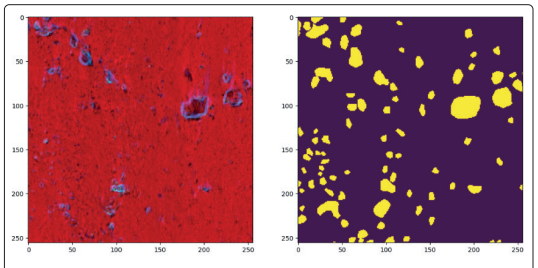

Fig. 3: Final composite image of slope, roughness and backscatter (left) with corresponding labels (right). The X and Y-axes are in pixel units

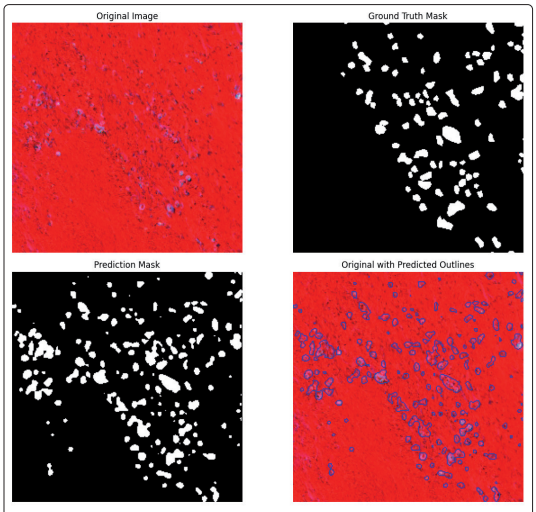

Fig. 4: Model-predicted rock boundaries overlaid on the original data, showing strong alignment with known features and some potential new detections

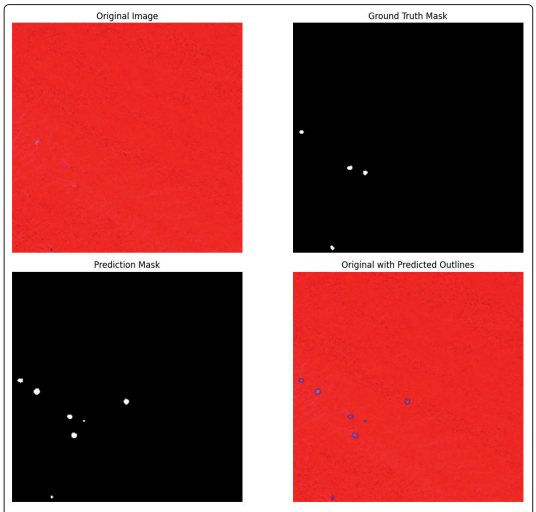

Fig. 5: Detection results in a low-density area, showing high precision and minimal over-prediction

identified as the most reliable configuration for rock detection in this study and was used for further detection.

Discussion

Interpretation of detection results

The model predictions show good spatial agreement with known rock formations, capturing key structural patterns. The integration of slope, roughness and backscatter data enabled the model to combine morphological and acoustic information, resulting in a richer representation of substrate characteristics. Detection confidence was highest in areas where rocks protruded clearly from surrounding sediments, highlighting the importance of elevation gradients and textural contrast. Across Fig. 4 and Fig. 5, the model consistently identified major rock features, particularly in densely populated zones, while also detecting smaller structures. Some of these additional detections may represent previously unmapped quartzite blocks, although others could be false positives caused by local texture variations or sedimentary features. In a few cases, the model segmented contiguous features differently from manual annotations, either splitting single units or merging adjacent ones, reflecting differences in boundary interpretation. In lower-density areas, the model maintained good detection performance while limiting false positives, indicating a relatively conservative response under low signal conditions.

Comparison with manual mapping

The model achieved a spatial accuracy comparable to manual mapping while offering substantially improved efficiency. To quantify performance, predicted rock areas were compared with manually delineated features using two representative test images. Analysis focused on overlapping regions to ensure direct comparability. Scatter plots and Bland-Altman analyses were used to evaluate agreement and assess potential bias in size estimation. Fig. 6 and Fig. 7 illustrate the segmentation results, showing overall strong agreement between predicted and manual masks, with localised discrepancies primarily along feature boundaries.

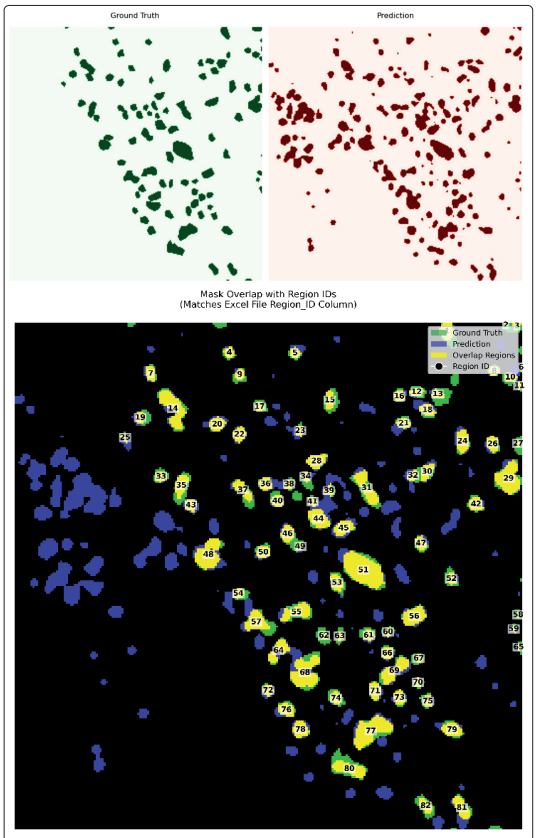

Fig. 6: Segmentation results for the first test image, showing manual ground truth (left), predicted mask (right), and colour-coded overlap (bottom)

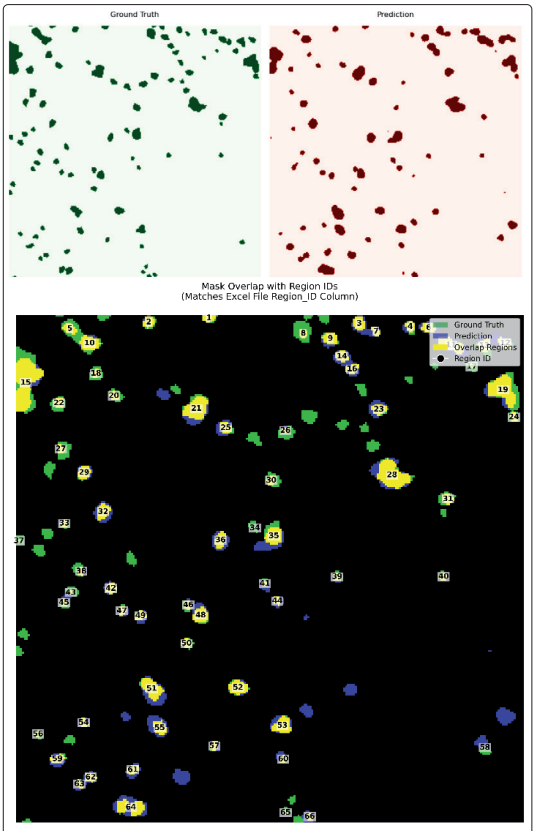

Fig. 7: Segmentation results for the second test image, showing manual ground truth (left), predicted mask (right), and colour-coded overlap (bottom)

Scatter plot analysis

Scatter plots (Fig. 8 and Fig. 9) show a strong linear relationship between predicted and ground truth rock areas, with coefficients of determination of $R^2 = 0.87$ and $R^2 = 0.80$, respectively. Most values cluster near the 1:1 line, indicating good agreement across a range of feature sizes. However, both cases show a tendency to underestimate larger rock areas, as indicated by the regression lines falling below the line of equality. This likely reflects smoothing effects or challenges in fully delineating complex, irregular geometries. Despite this limitation, the model performs reliably for small to medium-sized features.

Bland-Altman analysis

Bland-Altman plots (Fig. 10 and Fig. 11) further demonstrate reasonable agreement between predicted and manual measurements. In these plots, the average rock area (pixel^2) is shown on the X-axis, while the difference between predicted and manual rock area is shown on the Y-axis. For the first test image, a small negative bias (–4.90) indicates slight underestimation, with most values falling within the 95 % limits of agreement (–40.87 to +31.07). The second image shows a small positive bias (+3.15) and slightly wider limits of agreement (–46.61 to +52.91), reflecting greater variability. In both cases, variability increases with feature size, suggesting reduced accuracy for larger and more complex rock formations. A small number of outliers are present, likely due to segmentation challenges or irregular object shapes.

General interpretation

Overall, the CNN U-Net model demonstrates good agreement with manual mapping, as reflected by high R^2 values and relatively small biases. The absence of a consistent over- or underestimation trend suggests that errors are image-dependent and influenced by local conditions such as feature complexity and data quality. This variability indicates a degree of adaptability, which may be advantageous in heterogeneous fluvial environments. Manual mapping is time-intensive and subject to interpreter variability, particularly in ambiguous areas. In contrast, the model provides a consistent and reproducible approach to feature

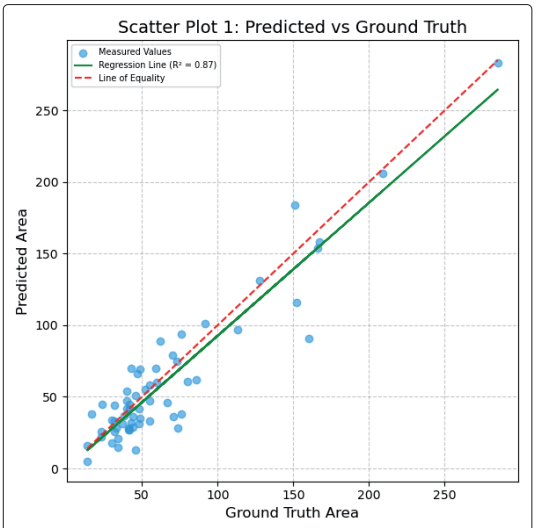

Fig. 8: Scatter plot of rock regions for the first test image showing area (pixel^2)

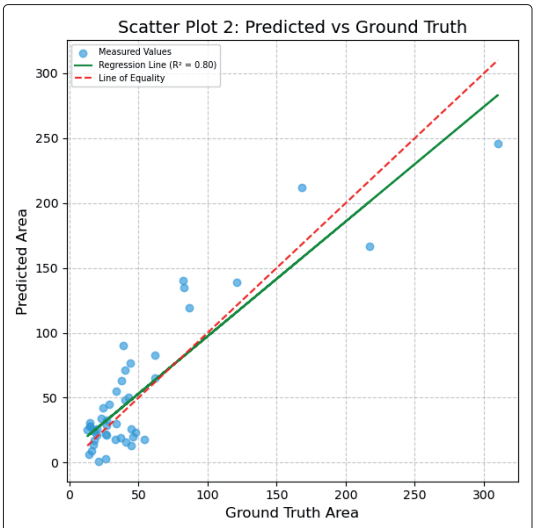

Fig. 9: Scatter plot of rock regions for the second test image showing area (pixel^2)

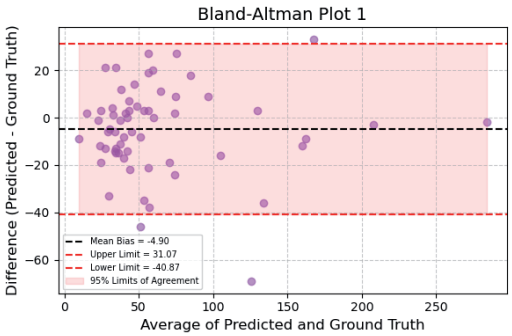

Fig. 10: Bland-Altman plot comparing ground truth and predicted rock areas for the first test image, with area expressed in pixel^2

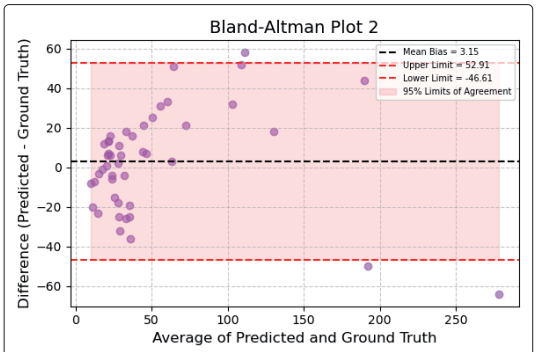

Fig. 11: Bland-Altman plot comparing ground truth and predicted rock areas for the second test image, with area expressed in pixel^2

detection. Some discrepancies between the two methods likely reflect both the subjectivity of manual interpretation and limitations in the training data. In addition to improving efficiency and scalability, it demonstrates the capability to detect subtle or previously unrecognised features, thereby complementing traditional mapping approaches.

Conclusion

The study highlights the value of integrating multi-source geophysical data, which significantly improved detection accuracy and robustness. The approach has clear potential to enhance hydrographic workflows by reducing reliance on manual interpretation, increasing consistency and enabling more efficient large-scale data processing. Despite these promising results, some limitations remain. The relatively small dataset and focus on a single study area may limit model generalisation. In addition, the use of non-expert annotations introduces potential label uncertainty, particularly for subtle or partially obscured features. Future work should prioritise expanding datasets with high-quality expert annotations and greater environmental diversity. Incorporating temporal data could also enable monitoring of riverbed dynamics over time. Overall, AI-driven approaches offer significant potential to improve navigational safety, optimise hydrographic survey workflows, and support sustainable management of inland waterways. //

References

Boenigk, Wolfgang (2021): Fluvial development of the Lower Rhine Basin during the late Tertiary and early Quaternary. DEUQUA Special Publications, DOI: 10.5194/deuquasp-3-125-2021

Cai, Lei; Jingyang Gao; Di Zhao (2020): A review of the application of deep learning in medical image classification and segmentation. Annals of Translational Medicine, DOI: 10.21037/atm.2020.02.44

Kraft, Markus; Nils Hollman; Kelly Torres et al. (2025): Detection of quartzite blocks in the River Rhine – Development of semi-automatic and automatic approaches for the detection of boulders. Journal of Applied Hydrography; DOI: 10.23784/HN130-04

LeCun, Yann; Yoshua Bengio; Geoffrey Hinton (2015): Deep learning. Nature, DOI: 10.1038/nature14539