Journal of Hydrology: Regional Studies 57 (2025) 102148

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Deep learning-enhanced detection of road culverts in high-resolution digital elevation models: Improving stream network accuracy in Sweden

William Lidberg

Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, Umeå 901 83, Sweden

A R T I C L E I N F O

Keywords:
Stream network
UNet
Machine learning
LiDAR
Culvert
ALS

A B S T R A C T

Study region: Sweden, a mostly forested country with many small forest roads obstructing topographical modelling of shallow groundwater and streams.

Study focus: Maps have traditionally been constructed from aerial photos, but dense forest canopies often obscure these streams from view. Topographical modelling is a widely adopted method for mapping small streams and has been proposed as a potential solution. However, road embankments can disrupt flow paths, acting like dams in the digital landscape. This study presents a novel method where a unique dataset of 28,512 culverts was used to develop a deep learning method to map road culverts, enabling the correction of digital elevation models and enhancing the accuracy of topographically derived stream networks.

New hydrological insights for the region: The deep learning model successfully mapped 87 % of all culverts in the test data, and integrating these predicted culverts into topographical models slightly improved the precision of stream networks extracted from high-resolution digital elevation models in the region. These findings demonstrate the possibility for UNet models to enhance hydrological modelling and stream network mapping but further research should focus on reducing the number of false positive culvert predictions.

1. Introduction

Headwater streams dominate surface water drainage networks in terms of stream length, and studies of headwater catchments have provided an understanding of different sources and controls on biogeochemistry in streams (Seibert et al., 2009; Tiwari et al., 2017), and also interactions with the atmosphere (Natchimuthu et al., 2017; Wallin et al., 2018). Intermittent and ephemeral streams represent half of the global river network (Schneider et al., 2017). Despite the importance of headwaters, it's a paradox that, at the same time, these streams are generally missing from stream databases (Bishop et al., 2008) and left without proper protection during forestry (Hafen et al., 2020). This is especially the case for intermittent and ephemeral streams (Fovet et al., 2021). One important reason for this is that small streams are poorly mapped (Jensen et al., 2017). Maps have traditionally been constructed from aerial photos, but small stream channels are difficult to observe from the air, especially under a dense tree canopy. Hence, there is a bias on maps that show larger streams and streams in agricultural areas where trees are scarce while smaller streams are missing (Benstead and Leigh, 2012; Skoulikidis et al., 2017).

To facilitate the protection of surface waters, the first step is to map streams so protection can be incorporated into everyday land-

E-mail address: William.lidberg@slu.se.

https://doi.org/10.1016/j.ejrh.2024.102148
Received 3 September 2024; Received in revised form 13 December 2024; Accepted 21 December 2024
Available online 9 January 2025

use planning and management. A growing number of national Airborne Laser Scanning (ALS) campaigns indicate a potential improvement in mapping accuracy and a positive step toward addressing these challenges (Barber and Shortridge, 2005; Murphy et al., 2008). Utilising high-resolution digital elevation models (DEM) obtained through airborne light detection and ranging (LiDAR) for topographical modelling is a widely adopted method for mapping small streams and has been proposed as a potential way to map streams missing from topographical maps (Jensen et al., 2017). This approach offers certain advantages, such as forming integrated drainage networks (Vaze and Teng, 2007) and adherence to stream channels in the DEM (Murphy et al., 2008). However, a DEM must undergo adjustments before any hydrological modelling to ensure hydrological accuracy (O'Callaghan and Mark, 1984). Removal of sinks is essential in allowing water to flow downhill towards the outlet, with sinks being defined as areas enclosed by cells with higher elevations, hindering water movement (O'Callaghan and Mark, 1984; Lindsay, 2015). Sinks can either be actual depressions in the landscape, but with very high resolution, DEM artefacts from urban features like bridges and road culverts can act like sinks. Lidberg et al. (2017) evaluated pre-processing methods in watersheds containing a large number of roads and concluded that higher resolution DEMs resulted in more accurate stream networks and that carving channels (breaching) across road embankments in the DEM reduced the impact on the modelled stream network. Pre-processing algorithms that breach a path across obstacles instead of filling up sinks are preferred but do not always breach the road embankment at the appropriate location (Lessard et al., 2023). Most of the world's streams are affected by dams, weirs and especially road culverts to some extent, and Mueller et al. (2011) demonstrated that these obstacles can alter community structure, productivity and the diversity of stream ecosystems. If the location of all culverts were known, it would be a simple task of carving paths across road embankments at the correct locations (Bryndal and Kroczak, 2019). Still, most culverts are old and not well-documented (Poppenga and Worstell, 2016). Using existing stream networks to burn channels across road embankments can be helpful (Lindsay, 2016a; Paz et al., 2008) but has limited effectiveness since small streams are poorly mapped. Melniks et al. (2020) used a mapped pre-existing ditch network to identify the possible culvert locations. Another approach is to use the road network to find possible locations of culverts (Wang and Fareed, 2021) or use road networks to completely remove roads from DEMs (Wang and Fareed, 2021; Van Nieuwenhuizen et al., 2021), but this can lead to problems if there is a discrepancy between the DEM and the road network data (Mäkinen et al., 2019). Additionally, road databases are not always up to date (Roussel et al., 2022), and small forest roads, in particular, are poorly mapped (White et al., 2010). Further, most road network databases don't include access roads to nearby agricultural fields since they are not part of the road network. Hence, it is imperative to create automated techniques for identifying culverts without pre-existing knowledge of road, ditch or stream locations.

A less explored method is to use machine learning to map culvert locations. Deep learning has been used to map bridges from satellite imagery (Pravalika et al., 2023; Qiu et al., 2023; Yang et al., 2021), Synthetic aperture radar (Chen et al., 2021) and ground-based LiDAR point clouds. Stanislawski et al. (2018) used a convolutional neural network (CNN) to map stream-road intersects in an agricultural area in central Iowa but this method requires a road network to operate. Wu et al. (2023) recently combined a CNN with an aerial imager and high-resolution DEM to map locations of drainage crossings in an agricultural area in Nebraska. The CNN model developed by Wu et al. (2023) had an accuracy of 90 % but is limited to classifying images as either drainage crossings or

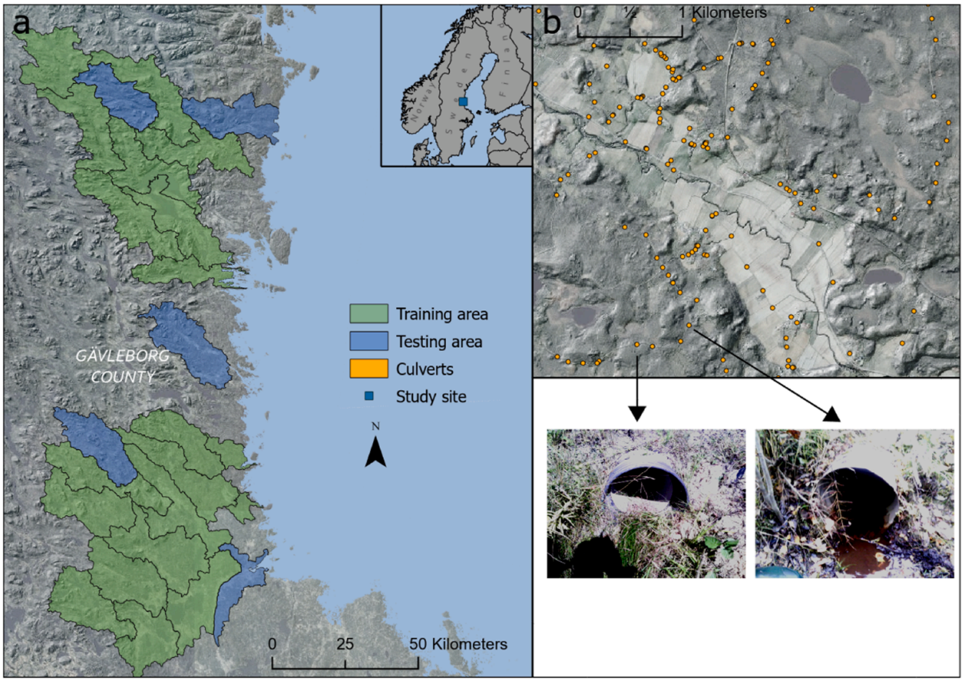

Fig. 1. Study areas with manually field-mapped road culverts. a) The study area consisted of 25 catchments, where 20 catchments (green) were used for training, and 5 catchments (blue) were used for testing. b) Example culvert locations within one of the test catchments.

non-drainage crossings. This makes it harder to use the model to breach the road at the right location for further processing. However, the UNet deep learning architecture introduced by (Ronneberger et al., 2015) gives each pixel in the input image a classification, resulting in a binary raster that can be subtracted from the DEM to carve paths across roads. Therefore, the research questions in this study were 1) to investigate whether road culverts could be automatically mapped using UNets 2) Whether this approach is practical in terms of processing time required to implement on national scales. 3) Do the accuracy of topographically derived stream networks improve when detected road culverts are included in the pre-processing of fine-resolution DEMs? 4) Are topographically derived streams more accurate than manually mapped streams on existing maps?

2. Method

This study uses data from two field campaigns carried out in Sweden and two national remote sensing datasets. The first field dataset was from a road culvert survey carried out in central Sweden and is used to train and evaluate deep-learning models to map culvert locations. The second field dataset was stream observations from the National Inventory of Landscapes in Sweden (NILS) (Ståhl et al., 2011) and is used to evaluate the accuracy of modelled stream networks. The two remote sensing datasets were from a national airborne laser scan and orthophotos.

2.1. Culvert survey

An extensive culvert survey was conducted in 25 watersheds in central Sweden by the Swedish Forest Agency during the snow-free periods of 2014–2017 (Fig. 1a). A total of 28,512 culverts were mapped with a handheld GPS with a horizontal accuracy of 0.3 m. Densely populated urban areas with underground drainage systems were excluded from the survey (0.3 % of the combined area). All other culverts are included. This includes culverts running alongside roads and underneath access roads to agricultural fields, access roads to forest stands, and driveways. The coordinates of both ends of each culvert were measured, and metrics such as diameter, length, material, working condition, and sediment accumulation were collected for most of the culverts. Additional metrics, such as the elevation difference between the outlet and stream water level, were manually measured with a ruler. The inventoried watersheds were split up into training and testing data, where 20 watersheds (23,304 culverts) were used for training, and five watersheds (5208 culverts) were used for testing. This means that 78 % of the culverts are used for training and 22 % for testing (Fig. 1b). The culverts were mainly small road culverts with an average diameter of 0.34 m and an average length of 6.7 m (Fig. 1c). Most of the culverts were constructed in steel at 61 %, followed by plastic at 29 % and concrete at 10 %. The average accumulation of sediment in the culverts covered 33 % of the cross-section area and 87 % of the culverts were considered migration obstacles for fish. The field-mapped culvert lines were buffered by 0.5 m and converted into a binary raster to be used as labels for training the deep learning model.

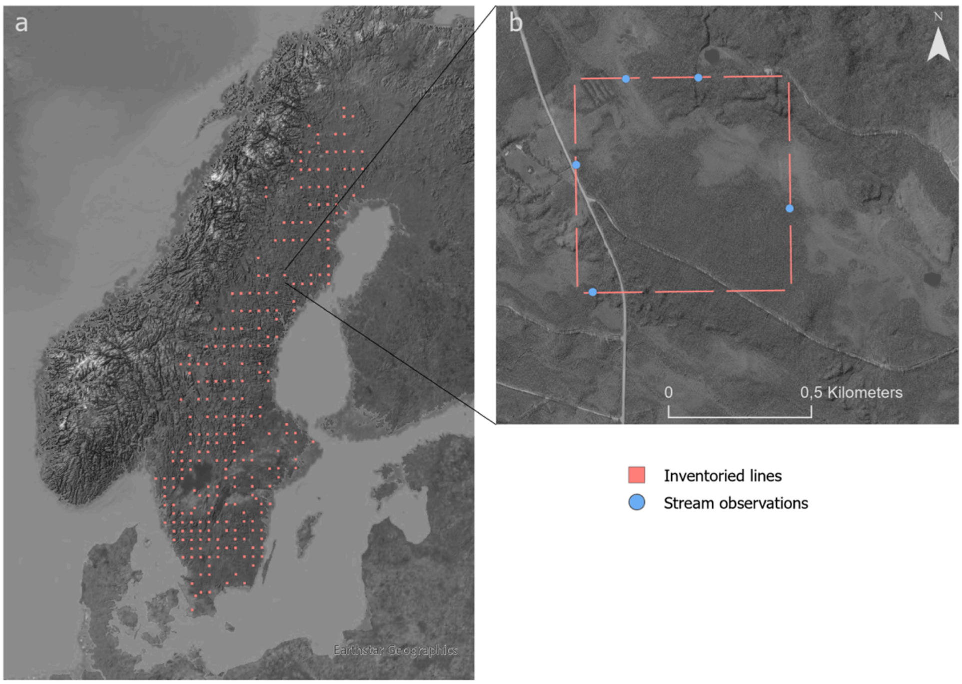

Fig. 2. Illustration of the NILS database. a) Locations of the line inventories across Sweden. b) Example of channel observations along one line inventory.

2.2. Stream observations

The NILS inventory encompasses 631 line inventories arranged in a square pattern (Fig. 2b) systematically distributed across Sweden, covering diverse landscapes such as forests, agricultural areas, mountains, wetlands, shores, and urban environments (Ståhl et al., 2011). This inventory was specifically devised to ensure statistically sound estimates for the entire country while also capturing the variability inherent in rare landscapes. 5, 364 channels were covered by ALS data and included in this study. Channels are defined as ditches or streams less than 6 m wide and the average channel width was 1.4 m meters.

2.3. Remote sensing data

A compact laser-based system (Leica ALS80-HP-8236) was used to collect the ALS data from an aircraft flying at 2888–3000 m. The ALS point clouds had a point density of 1–2 points m^{-2} and a DEM with 0.5 m resolution was created from the ALS point clouds using a TIN gridding approach implemented in Whitebox tools 2.2.0 (Lindsay, 2016b). While it is not generally recommended to create such higher resolution DEM from point clouds with only 1–2 points per square meter, previous research using this point cloud showed that 0.5 m outperforms 1 m for detecting hunting pits (Lidberg et al., 2024) and ditches (Laudon et al., 2022; Lidberg et al., 2023). Additionally, the orthophotos had a resolution of 0.5 m. Therefore a DEM with matching resolution of 0.5 m was created instead of downsampling the orthophotos.

The DEM was normalised by calculating a topographical index instead of using the raw elevation values. The idea was to highlight local topography, making it easier for the deep-learning model to learn how culvert locations appear in the DEMs. Lidberg et al. (2023) evaluated a range of different topographical indices for detecting hunting pits in a high-resolution DEM and found that the choice of the topographical index did not have a major impact on the deep learning models' ability to extract information. The topographical index "max downslope elevation change" was selected as a method to normalise the DEM without the requirement to specify any parameters to the algorithm. Max downslope elevation change represents the maximum elevation drop between each grid cell and its neighbouring cells. This typically resulted in values between 0 and 10 in the study area, and In addition to the topographical data, I included orthophotos from aerial imagery. The orthophotos had three bands (red, green and blue) in 8-bit colour depth and had a resolution of 0.5 m. The LiDAR data and orthophotos were downloaded from the Swedish mapping, cadastral and land registration authority.

2.4. Deep learning architectures

This study evaluated the ability of two deep-learning architectures to identify road culverts from the remote sensing data. The two architectures were a standard UNet introduced by (Ronneberger et al., 2015) and a UNet with both Residual connections and Attention blocks. TensorFlow 2.16 was used to build the two types of encoder-decoder-style deep neural networks to transform the topographical indices into images highlighting the location of road culverts. On the encoding path, the UNet learn a series of filters, organised in convolutional blocks, which express larger and larger neighbourhoods of pixels in fewer and fewer vectors of features. This down-sampling forces the networks to ignore noise and extract features relevant to culvert detection. In addition to this regular UNet, I built a residual attention UNet (RA-UNet), where the convolutional blocks were replaced with residual convolutional blocks that continuously update the calculated gradient values in each convolutional block to combat the vanishing gradient problem (Ni et al., 2019). Finally, attention blocks were added to the decoder. The attention mechanism in the UNet architecture was introduced by (Oktay et al., 2018) for medical imaging and the idea was to give relevant parts of the image bigger weights than the less relevant parts. Ghaznavi et al. (2022) found that applying the residual and attention mechanisms together resulted in the most accurate results on cell segmentation in microscopy images. The models were trained on three different combinations of the remote sensing data. First, the models were trained on the topographical index Max downslope elevation change, then on the orthophotos and finally, a combination of the topographical index and orthophotos.

Since only about 1 % of the pixels in the input images were labelled as culverts, I used a focal loss function to increase their weight during training (Lin et al., 2017). Further, I assumed that the culverts and the immediate surrounding terrain had no preferred orientation, allowing for image augmentation by rotating image chips during training. All models were trained for 50 epochs using a batch size of 32. Training and inference were done using an NVIDIA RTX A6000 GPU and AMD Ryzen Threadripper 3990X Processor. The performance of various deep learning architectures was evaluated using culvert masks from five test catchments. The evaluation metrics included true positives, false positives, true negatives, false negatives, and the Matthews correlation coefficient (MCC). The MCC is a statistical measure that yields a high score only when the model performs well across all four confusion matrix categories (true positives, false negatives, true negatives, and false positives) (Chicco and Jurman, 2020). The model with the highest MCC was selected to correct the digital elevation model (DEM) for topographical modeling of stream networks.

To assess how accurately the model predicts culvert sizes, labeled masks were compared with predicted culvert pixels. This is particularly relevant when incorporating culverts into a DEM. Since culverts span multiple pixels, the detected culvert pixels were also converted into polygons and compared to field-mapped culverts through spatial intersection. A predicted culvert was considered a true positive only if part of its polygon directly intersected with a field-mapped culvert.

The predicted culverts were then used to pre-process a 0.5-meter resolution DEM in the test catchments and a 1-meter DEM of Sweden for further evaluation.

2.5. Topographical stream modelling

2.5.1. Evaluation with culverts

Topographically extracted stream networks were compared to inlets and outlets of culvert locations in the test catchments in order to answer the second research question regarding the impact of detected culverts on stream modelling. The 0.5 m DEMs of the test catchments were resampled to 1 m due to the low point density of the ALS data. Filling has been a common way to pre-process DEMs for stream extractions, so the fill algorithm by (Wang and Liu, 2006) was used to set a baseline. However, Lidberg et al. (2017) argued that roads are best handled using breaching, so I also evaluated stream networks preprocessed using the least-cost breaching approach described by (Lindsay and Dhun, 2015). Another approach is to remove roads entirely, as suggested by Van Nieuwenhuizen et al. (2021), so this method was also evaluated. Finally, the DEM was pre-processed by burning the culverts predicted by the machine learning model into the DEM before the same least-cost breaching was applied (Fig. 3). Flow pointer and flow accumulation were then calculated from the pre-processed DEMs using Deterministic 8 (O'Callaghan and Mark, 1984), and streams were extracted using a stream initiation threshold (representing the catchment area required to form a stream) of 10,000 m^2 (1 ha) (Fig. 3). This is an intentionally low threshold aimed at capturing headwater streams during high-flow conditions in the region (Ågren et al., 2015). Most culverts are expected to have a very low flow accumulation as they are meant to accommodate seasonal streams and water from rain events. The extracted raster streams were converted to vector lines and evaluated with the field-mapped road culverts. A stream segment was considered correct if it intersected within 5 m of a culvert's inlet and outlet. The stream had to intersect within 5 m of both the inlet and outlet of the culvert for it to be considered a correct intersection. The 5-meter distance was used to determine whether a stream crossed the road at the culvert location or merely intersected the culvert's inlet before continuing parallel to the road without crossing. In this region, most roads are relatively narrow, so using a larger search distance would cause the stream to appear as though it intersects both the inlet and outlet of the culvert, even if it does not cross the road. The number of correct culvert intersections was used to estimate the accuracy of the road crossings from the respective stream networks in the test watersheds.

2.5.2. Evaluation with stream observations

The final research question was whether topographically derived stream networks are more accurate than manually mapped streams on existing maps, and it was answered using the NILS inventory of stream locations in Sweden. The stream observations from the NILS inventory were first intersected with subbasins from the Swedish Meteorological and hydrological institute. Resulting in an average of 10 observations within each watershed. The selected watersheds were then buffered by 2 km to avoid potential edge effects. A digital elevation model with a resolution of 1 m was created from the ALS point cloud for each watershed to evaluate the predicted impact of culverts on stream extraction. The pre-processing of the DEM was done by burning culverts predicted by the machine learning model into the DEM before the breaching algorithm was applied. D8-Flow accumulation values were then calculated to the line inventories used in the NILS dataset. Due to uncertainties in GPS positioning, certain field-inventoried stream channels were found to deviate by up to 30 m from their actual locations, particularly beneath dense forest canopies. To address this issue, the inventoried stream channel points were adjusted by snapping them to the highest flow accumulation along the line inventory within a 30-meter, thus reducing the effects of GPS inaccuracies. Multiple stream initiation thresholds from 1 ha to 50 ha were used to extract stream networks. The extracted streams were then evaluated against the stream observations in the NILS dataset. Stream networks from the

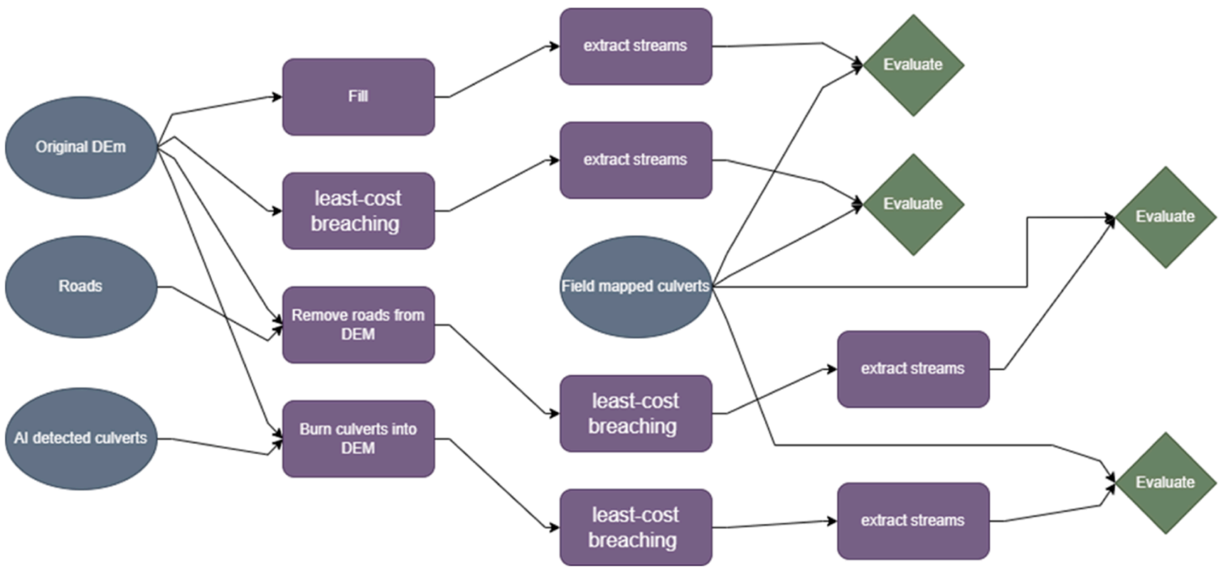

Fig. 3. Flowchart of the evaluation of topographically modelled streams and field mapped road culverts. The stream lines had to pass within 5 m of both the inlet and outlet of a field mapped road culvert to count as a correct intersection.

1:12 500 scale property map, produced by the Swedish Mapping, Cadastral and Land Registration Authority, were also evaluated to set a baseline to compare with the topographically modelled streams. Finally, since it is unreasonable to expect one fixed threshold to work well everywhere in Sweden, I also identified the optimal stream initiation threshold for each subbasin based on the most accurate stream network compared to stream channel observations in the NILS inventory. This was done in an iterative manner by looping over stream initiation thresholds from 1 ha to 50 ha and comparing each resulting steam network with observed stream channels. Finally, the optimal stream initiation threshold within each subbasin was used to create an optimal stream network.

3. Results

3.1. Mapping culverts

The RA-UNet trained on a combination of topographical data and orthophotos was the best method for mapping culverts. When evaluated against images and labels in the test data, it had an MCC value of 0.58, a precision of 47 %, and a recall of 74 %. Training exclusively on topographical data worked reasonably well, with an MCC of 0.55, but the model trained on orthophotos only had an MCC of 0.3 (Table 1).

Evaluating the predicted culvert pixels against the labelled culvert masks so the detected culvert pixels from the best model (RA-UNet combined data) were also converted to polygons and compared to the manually digitised culverts in the test watersheds (Fig. 4). The culverts predicted by the RA-UNet model intersected 4513 of the 5208 manually mapped culverts in the test watersheds (87 %) but also produced 19, 318 false positive culvert locations, resulting in a precision of 19 %. A false positive is when the model predicted a culvert that did exist in the field. True positive meant that the field inventoried culverts and the predicted culverts overlapped. A false negative is when the model fails to predict a field-inventoried culvert. The difference in recall and precision between the pixel-based evaluation and the evaluation of polygons comes from the fact that a culvert consists of multiple pixels.

When the deep learning model was applied to parts of Sweden with ALS coverage, it detected 4.9 million culverts. Assuming the precision is 19 % for the entire area, the number of real culverts would be 0.88 million (Fig. 5).

3.2. Topographical stream modelling

Streams extracted from filling the DEM correctly intersected both ends of 1155 culverts while least cost breaching intersected both ends of 1851 culverts. Removing roads entirely before the least cost breaching resulted in only 1388 intersections of both ends. Burning detected culverts into the DEM before least cost breaching was the most accurate and resulted in 1972 intersections. Burning culverts before least cost breaching improved the result near roads slightly while keeping the stream network intact upstream of roads (Fig. 6).

The method where detected culverts were burned into the DEM before breaching was also implemented nationally in Sweden, and the resulting stream network was compared to stream observations from the NILS database (Fig. 2). Stream initiation thresholds ranging from 1 ha to 50 ha were used to delineate the stream networks. Topographically modelled stream networks with stream initiation thresholds between 1 and 20 ha were more accurate than the current stream network on topographical maps (Fig. 7). However, MCC values decreased with stream initiation thresholds larger than 20 ha and continued to decline with increasing thresholds. The MCC score was 0.42 when the optimal stream initiation threshold was used for each subbasin.

4. Discussion

The RA-UNet model outperformed the standard UNet model in mapping culverts, regardless of whether the models were trained on topography, orthophotos, or a combination of both. This aligns with previous research where RA-UNets have shown superior performance on medical imagery (Li et al., 2022) and, more recently, on remote sensing applications such as mapping water bodies in the Czech Republic (Ghaznavi et al., 2024). While other studies have successfully mapped large features like bridges using deep learning combined with satellite imagery (Pravalika et al., 2023; Qiu et al., 2023; Yang et al., 2021) and Synthetic aperture radar (Chen et al., 2021), the smaller size and limited visibility of road culverts on satellite imagery have presented unique challenges.

The results presented here are helpful given that road embankments, bridges, and culverts are critical factors for creating accurate stream networks from fine-resolution ALS data (Schwanghart et al., 2013). The removal of 'artificial dams' caused by unaccounted road culverts is essential for ensuring that flow paths reflect actual surface drainage conditions (Bryndal and Kroczak, 2019). Although

Table 1
Summary of the different models when trained on combinations of topographical data and orthophotos. These metrics were based on the classified pixels from images in the test data. Recall measures the proportion of actual culvert pixels that the model correctly detected. Precision measures the proportion of predicted culvert pixels that are correctly labelled as culvert pixels.

	Recall	Precision	MCC
UNet topographical data	62 %	50 %	0.55
UNet orthophoto data	19 %	22 %	0.19
UNet combined data	66 %	46 %	0.54
RA-UNet topographical data	78 %	41 %	0.55
RA-UNet orthophoto data	37 %	25 %	0.3
RA-UNet combined data	74 %	47 %	**0.58**

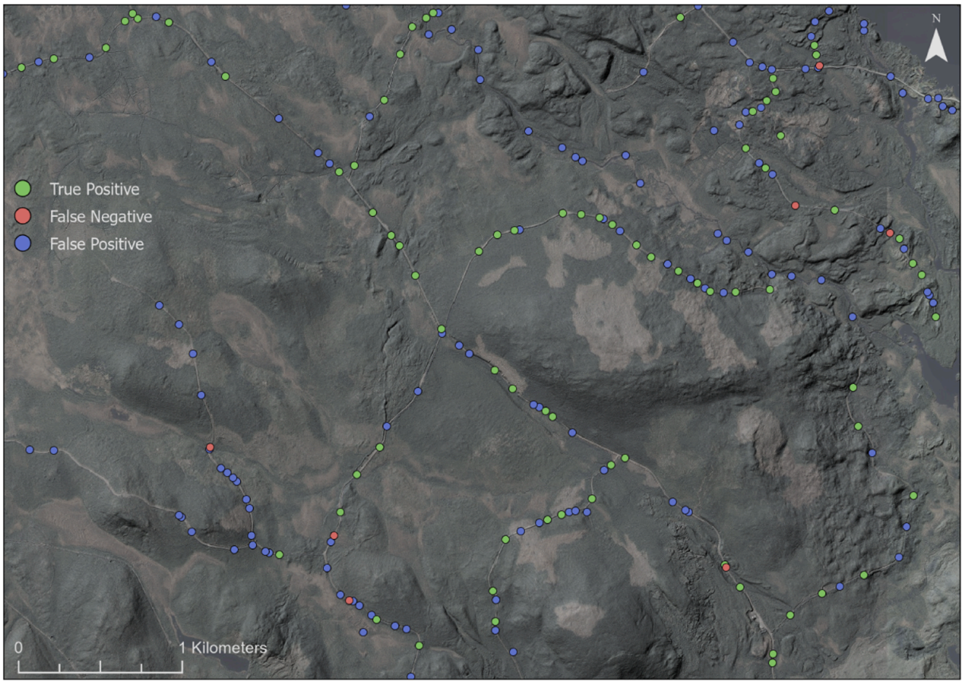

Fig. 4. Predicted culverts in a small section of the test watersheds. The model captured 87 % of all culverts in the test areas (green points). However, the model also produced many false positive predictions (purple points) but only a few false negatives (red points).

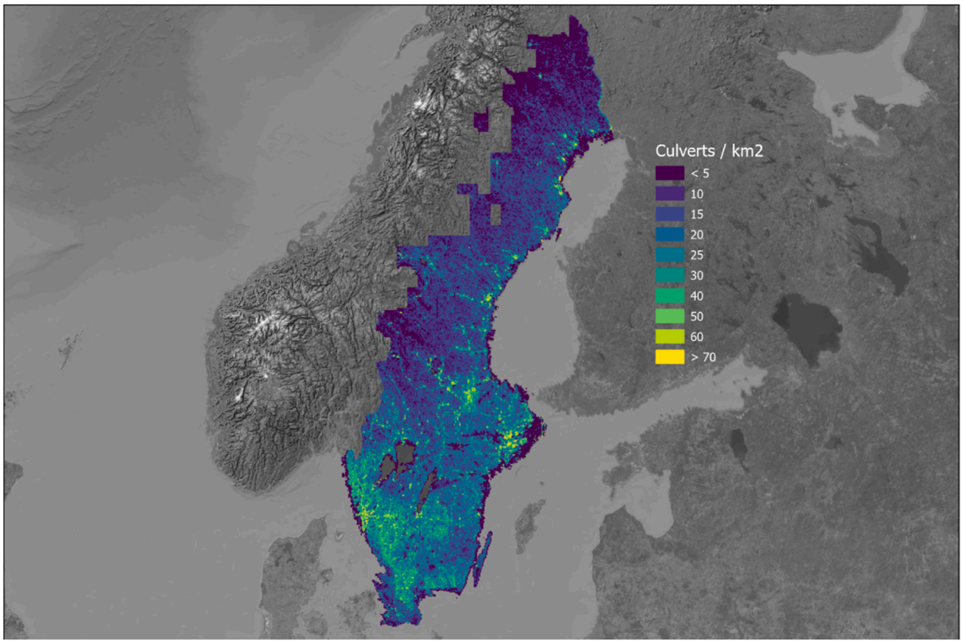

Fig. 5. Culvert density in the areas of Sweden covered by the ALS data. North-western Sweden was missing ALS data during this study and could not be included. The colours represent the number of detected culverts per square kilometre. Densely populated areas like Gothenburg and Stockholm have a higher culvert density than the more scarcely populated areas in the north.

reducing DEM resolution could mitigate some issues related to road banks, (Lidberg et al., 2017) found that higher-resolution DEMs (up to 2 m) produced more accurate stream networks, underscoring the necessity of accurately mapping road culverts.

There have been a number of traditional algorithms that aim to map road culverts from DEMs. Roelens et al. (2018) handled road-ditch intersections by connecting endpoints of ditches extracted from local small, low-relief features from DEMs. Melniks et al.

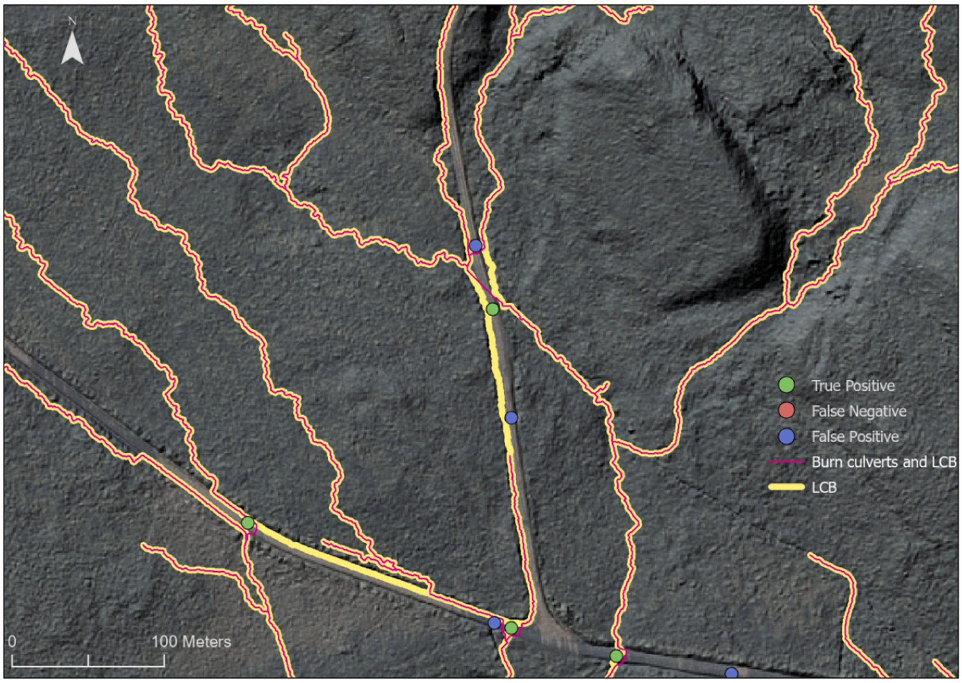

Fig. 6. Example of the two best stream networks, least cost breaching (LCB) and burning AI-detected culverts before least cost breaching. The LCB method is displayed as a ticker line to visualise the overlapping sections of the stream networks. The LCB method runs along the road instead of across it in the center of the image and in the bottom left.

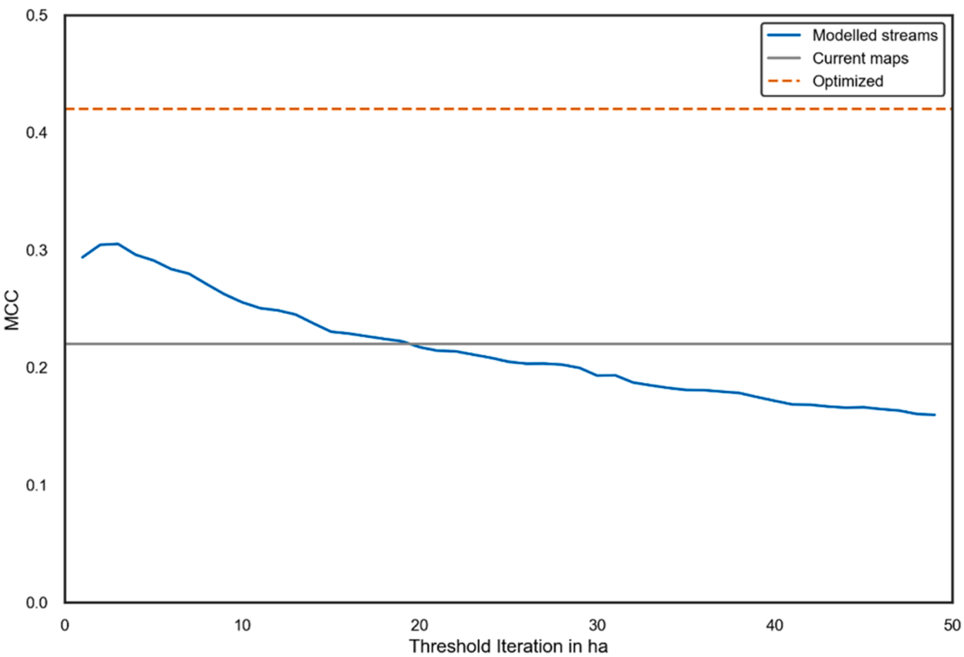

Fig. 7. Performance of the different pre-processing methods on stream network delineated compared to observed streams in the NILS database. Stream initiation thresholds from 1 to 50 were evaluated in 1-ha increments. The orange dashed line represents the MCC when optimal stream initiation thresholds were selected for each sub-basin.

(2020), suggested a similar approach where pre-existing ditch maps in combination with ALS data were used to map 37 % of the culverts in their study site. Mäkinen et al. (2019) modified the carving algorithm by (Soille et al., 2003) into culvert-aware carving and mapped 67 % of the culverts in their test area. However, these methods often rely on pre-existing ditch maps or specific parameter settings, which limit their transferability to new areas without preliminary testing. The deep learning approach presented in this study has demonstrated the ability to learn patterns from the data without the need for auxiliary data such as ditch networks or road databases. Additionally, this approach can identify culverts running alongside roads and underneath access roads to agricultural fields and forest stands, providing broader applicability.

While most culverts were detected (87 %), the deep learning model also suffered from a high false detection rate. This can be adjusted by different class weights used in the loss function during training, which would also result in a lower recall. The loss function “Binary Focal Crossentropy” was used here with class weights using alpha 0.97 and gamma 0.2. Conversely, the weights can be set to favour a lower precision but at the cost of lower recall. However, one of the reasons for mapping these culverts is to improve stream networks extracted from fine-resolution DEMs which motivated the focus on a higher recall. Future research could involve extracting additional topographical features, such as stream order, flow accumulation, distance to water divides, and flow path length to each predicted culvert. These attributes could be used to train a decision tree-based model to help filter out false positive culverts. A decision tree may capture logical relationships that convolutional filters in a UNet model cannot capture from the orthophotos and topography alone. One advantage of UNet over CNNs is that a UNet prouces a raster with the predicted culverts. This raster can be burned into the DEM which is more useful than classifying an entire image into a culvert or none culvert location. Another approach could be to combine a UNet as presented here, with a CNN like the one described by Wu et al. (2023) where a image is generated from the raster data centered around the culvert polygons predicted by the UNet model. This could be a method to post process the data to filter out false positives while keeping the advantage of the UNet raster output.

Increasing DEM resolution also increases the processing time, and mapping culverts using deep learning further increases the processing time. However, due to recent developments in advanced geospatial data analysis platform (Lindsay 2016b), this is no longer a major hurdle to overcome. Implementing deep learning models on a combination of 0.5 m DEM data and orthophotos is still very demanding in terms of the processing power required but it is still a feasible solution with the right hardware. The model trained without orthophotos only had a slightly lower MCC of 0.55 compared to 0. 58 with the orthophotos included, which means that users can rely on topographical indices without losing much in terms of accuracy. The deep learning model developed in this study only took five days to predict culverts in Sweden (Fig. 4) on a 40 GB Multi-Instance GPU from an NVIDIA A100. If the full card would be utilised for inference it would only take around a day to predict culverts for all of Sweden. Excluding data handling and pre-processing. Combining orthophotos with a topographical index also introduces some overhead to the process, and while the model can run on only DEM data, it did improve from MCC 0.55 to MCC 0.58 when the DEM was combined with orthophotos. This showed that there was some useful information in the RGB imagery that could be utilised to improve topographical stream models but only using a DEM also produced useful predictions. Future studies could explore incorporating additional data types, such as infrared imagery or time-series remote sensing data, to further enhance model accuracy. Another topic worth exploring could be whether the deep learning model performs differently depending on the size of the road or the size of the stream channel.

The least-cost breaching approach described by Lindsay and Dhun (2015) produced more accurate culvert intersections than traditional filling, which aligns with previous research (Lessard et al., 2023; Lidberg et al., 2017). However, integrating detected culverts into the DEM before breaching resulted in even more accurate stream-road intersections, demonstrating the possibility of the UNet model for practical hydrological applications. Finally, the study highlighted the importance of selecting appropriate stream initiation thresholds for accurate stream network delineation. The optimal thresholds varied across watersheds, but when chosen correctly, the modelled stream networks were more accurate than existing topographical maps, achieving an MCC of 0.44 compared to existing topographical maps which had MCC of 0.31. This can be done to some extent with the help of national inventories such as NILS and data on runoff, soil texture and local topography. Ultimately, more accurately modelled topographical stream networks would allow us to shed some light on the unknown headwaters described by (Christensen et al., 2022; Bishop et al., 2008). Forest managers could use these stream networks to better plan operations in wet areas near streams in order to prevent rutting (Lidberg et al., 2020) and subsequent sediment transport (Hasselquist et al., 2024), mercury export (Munthe and Hultberg, 2004), modelling surface drainage system during heavy rainfall (Bryndal et al., 2020) or design buffer zones during forestry operations (Chellaiah and Kuglerová, 2021; Hasselquist et al., 2024).

5. Conclusions

This study demonstrates the possibility of using deep learning to map road culverts and improve the accuracy of topographically derived stream networks. By incorporating detected culverts into DEMs before applying hydrological pre-processing, the accuracy of extracted stream networks, addressing the challenges posed by road embankments acting as dams in the digital landscape. The RA-UNet architecture, trained on a combination of topographical data and orthophotos, successfully identified 87 % of culverts in test areas, proving the model’s robustness and utility on a large scale. Future research could explore reducing the number of false positives and integrating additional data sources, such as infrared imagery or temporal remote sensing data.

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 article.

Acknowledgements

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program—Humanities and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation, the Marcus and Amalia Wallenberg Foundation and Kempestiftelserna. I would also like to thank the personnel at the Swedish Forest Agency who conducted the field survey for the five large study catchments and KUA Foundation Gävleborg; County Administrative Board in Gävleborg; Arbetsförmedlingen in Gävleborg; and the Swedish Mapping, Cadastral and Land Registration Authority in Gävleborg.

Data availability

Both data and code are available at https://doi.org/10.5878/rjpg-ec44

References

Ågren, A.M., Lidberg, W., Ring, E., 2015. Mapping temporal dynamics in a forest stream network—implications for riparian forest management. Forests 6 (9), 9. https://doi.org/10.3390/f6092982.

Barber, C.P., Shortridge, A., 2005. Lidar elevation data for surface hydrologic modeling: resolution and representation issues. Cartogr. Geogr. Inf. Sci. 32 (4), 401–410. https://doi.org/10.1559/152304005775194692.

Benstead, J.P., Leigh, D.S., 2012. An expanded role for river networks. Nat. Geosci. 5 (10), 678–679. https://doi.org/10.1038/ngeo1593.

Bishop, K., Buffam, I., Erlandsson, M., Folster, J., Laudon, H., Seibert, J., Temnerud, J., Fölster, J., Laudon, H., Seibert, J., Temnerud, J., 2008. Aqua Incognita: the unknown headwaters. Hydrol. Process. 22 (8), 1239–1242. https://doi.org/10.1002/hyp.7049.

Bryndal, T., Kroczak, R., 2019. Reconstruction and characterization of the surface drainage system functioning during extreme rainfall: the analysis with use of the ALS-LIDAR data—the case study in two small flysch catchments (Outer Carpathian, Poland). Environ. Earth Sci. 78 (6), 215. https://doi.org/10.1007/s12665-019-8211-6.

Bryndal, T., Kroczak, R., Kijowska-Strugała, M., Bochenek, W., 2020. How human interference changes the drainage network operating during heavy rainfalls in a medium-high relief flysch mountain catchment? The case study of the Bystrzanka catchment (Outer Carpathians, Poland). CATENA 194, 104662. https://doi.org/10.1016/j.catena.2020.104662.

Chellaiah, D., Kuglerová, L., 2021. Are riparian buffers surrounding forestry-impacted streams sufficient to meet key ecological objectives? A Swedish case study. For. Ecol. Manag. 499, 119591. https://doi.org/10.1016/j.foreco.2021.119591.

Chen, L., Weng, T., Xing, J., Li, Z., Yuan, Z., Pan, Z., Tan, S., Luo, R., 2021. Employing deep learning for automatic river bridge detection from SAR images based on Adaptively effective feature fusion. Int. J. Appl. Earth Obs. Geoinf. 102, 102425. https://doi.org/10.1016/j.jag.2021.102425.

Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genom. 21 (1), 6. https://doi.org/10.1186/s12864-019-6413-7.

Christensen, J.R., Golden, H.E., Alexander, L.C., Pickard, B.R., Fritz, K.M., Lane, C.R., Weber, M.H., Kwok, R.M., Keefer, M.N., 2022. Headwater streams and inland wetlands: status and advancements of geospatial datasets and maps across the United States. Earth-Sci. Rev. 235, 104230. https://doi.org/10.1016/j.earscirev.2022.104230.

Fovet, O., Belemtougri, A., Boithias, L., Braud, I., Charlier, J.-B., Cottet, M., Daudin, K., Dramais, G., Ducharne, A., Folton, N., Grippa, M., Hector, B., Kuppel, S., Le Coz, J., Legal, L., Martin, P., Moatar, F., Molénat, J., Probst, A., Datry, T., 2021. Intermittent rivers and ephemeral streams: perspectives for critical zone science and research on socio-ecosystems. WIREs Water 8 (4), e1523. https://doi.org/10.1002/wat2.1523.

Ghaznavi, A., Rychtáriková, R., Saberioon, M., Štys, D., 2022. Cell segmentation from telecentric bright-field transmitted light microscopy images using a Residual Attention U-Net: a case study on HeLa line. Comput. Biol. Med. 147, 105805. https://doi.org/10.1016/j.compbiomed.2022.105805.

Ghaznavi, A., Saberioon, M., Brom, J., Itzerott, S., 2024. Comparative performance analysis of simple U-Net, residual attention U-Net, and VGG16-U-Net for inventory inland water bodies. Appl. Comput. Geosci. 21, 100150. https://doi.org/10.1016/j.acags.2023.100150.

Hafen, K.C., Blasch, K.W., Rea, A., Sando, R., Gessler, P.E., 2020. The influence of climate variability on the accuracy of NHD perennial and nonperennial stream classifications. JAWRA J. Am. Water Resour. Assoc. 56 (5), 903–916. https://doi.org/10.1111/1752-1688.12871.

Hasselquist, E.M., Polvi, L.E., Staaf, R., Winkowska, M., Baan Hofman, R., Kuglerová, L., 2024. The role of riparian buffer width on sediment connectivity through windthrow in a boreal headwater stream. Geomorphology 461, 109320. https://doi.org/10.1016/j.geomorph.2024.109320.

Jensen, C.K., McGuire, K.J., Prince, P.S., 2017. Headwater stream length dynamics across four physiographic provinces of the Appalachian Highlands. Hydrol. Process. 31 (19), 3350–3363. https://doi.org/10.1002/hyp.11259.

Laudon, H., Lidberg, W., Sponseller, R.A., Maher Hasselquist, E., Westphal, F., Östlund, L., Sandström, C., Järveoja, J., Peichl, M., Ågren, A.M., 2022. Emerging technology can guide ecosystem restoration for future water security. Hydrol. Process. 36 (10), e14729. https://doi.org/10.1002/hyp.14729.

Lessard, F., Jutras, S., Perreault, N., Guilbert, É., 2023. Performance of automated geoprocessing methods for culvert detection in remote Forest environments. Can. Water Resour. J. / Rev. Can. Des. Ressour. Hydr. 48 (3), 248–257. https://doi.org/10.1080/07011784.2022.2160660.

Li, Z., Zhang, H., Li, Z., Ren, Z., 2022. Residual-attention UNet++: a nested residual-attention U-net for medical image segmentation. Appl. Sci. 12 (14), 14. https://doi.org/10.3390/app12147149.

Lidberg, W., Nilsson, M., Lundmark, T., Ågren, A.M., 2017. Evaluating preprocessing methods of digital elevation models for hydrological modelling. Hydrol. Process. https://doi.org/10.1002/hyp.11385.

Lidberg, W., Nilsson, M., Ågren, A., 2020. Using machine learning to generate high-resolution wet area maps for planning forest management: a study in a boreal forest landscape. Ambio 49 (2). https://doi.org/10.1007/s13280-019-01196-9.

Lidberg, W., Paul, S.S., Westphal, F., Richter, K.F., Lavesson, N., Melniks, R., Ivanovs, J., Ciesielski, M., Leinonen, A., Ågren, A.M., 2023. Mapping drainage ditches in forested landscapes using deep learning and aerial laser scanning. J. Irrig. Drain. Eng. 149 (3), 04022051. https://doi.org/10.1061/JIDEDH.IRENG-9796.

Lidberg, W., Westphal, F., Brax, C., Sandström, C., Östlund, L., 2024. Detection of hunting pits using airborne laser scanning and deep learning. J. Field Archaeol. 49 (6), 395–405. https://doi.org/10.1080/00934690.2024.2364428.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). *Focal Loss for Dense Object Detection*. arXiv.Org. ⟨https://arxiv.org/abs/1708.02002v2⟩.

Lindsay, J.B., 2016b. Whitebox GAT: a case study in geomorphometric analysis. Comput. Geosci. 95, 75–84. https://doi.org/10.1016/j.cageo.2016.07.003.

Lindsay, J.B., 2016a. The practice of DEM stream burning revisited. Earth Surf. Process. Landf. 41 (5), 658–668. https://doi.org/10.1002/esp.3888.

Lindsay, J.B., Dhun, K., 2015. Modelling surface drainage patterns in altered landscapes using LiDAR. Int. J. Geogr. Inf. Sci. 29 (3), 397–411.

Mäkinen, V., Oksanen, J., Sarjakoski, T., 2019. Automatic determination of stream networks from DEMs by using road network data to locate culverts. Int. J. Geogr. Inf. Sci. 33 (2), 291–313. https://doi.org/10.1080/13658816.2018.1530353.

Melniks, R., Ivanovs, J., Lazdins, A., 2020. Identification of possible ditch culvert locations using lidar data. Eng. Rural Dev. 19, 1706–1711. https://doi.org/10.22616/ERDev.2020.19.TF439 (Scopus).

Mueller, M., Pander, J., Geist, J., 2011. The effects of weirs on structural stream habitat and biological communities. J. Appl. Ecol. 48 (6), 1450–1461. https://doi.org/10.1111/j.1365-2664.2011.02035.x.

Munthe, J., Hultberg, H., 2004. Mercury and methylmercury in runoff from a forested catchment—Concentrations, fluxes, and their response to manipulations. Water Air Soil Pollut.: Focus 4 (2–3), 607–618. https://doi.org/10.1023/B:WAFO.0000028381.04393.ed.

Murphy, P., Ogilvie, J., Meng, F., Arp, P., 2008. Stream network modelling using lidar and photogrammetric digital elevation models: a comparison and field verification. Hydrol. Process. 22 (August 2007), 1747–1754. https://doi.org/10.1002/hyp.

Natchimuthu, S., Wallin, M.B., Klemedtsson, L., Bastviken, D., 2017. Spatio-temporal patterns of stream methane and carbon dioxide emissions in a hemiboreal catchment in Southwest Sweden. Sci. Rep. 7 (April 2016). https://doi.org/10.1038/srep39729, 39729–39729.

Ni, Z.-L., Bian, G.-B., Zhou, X.-H., Hou, Z.-G., Xie, X.-L., Wang, C., Zhou, Y.-J., Li, R.-Q., & Li, Z. (2019). *RAUNet: Residual Attention U-Net for Semantic Segmentation of Cataract Surgical Instruments* (arXiv:1909.10360). arXiv. ⟨https://doi.org/10.48550/arXiv.1909.10360⟩.

O'Callaghan, J.F., Mark, D.M., 1984. The extraction of drainage networks from digital elevation data. Comput. Vis. Graph. Image Process. 28 (3), 323–344. https://doi.org/10.1016/S0734-189X(84)80011-0.

Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., & Rueckert, D. (2018). *Attention U-Net: Learning Where to Look for the Pancreas* (arXiv:1804.03999). arXiv. ⟨https://doi.org/10.48550/arXiv.1804.03999⟩.

Paz, A.R., da, Collischonn, W., Risso, A., Mendes, C.A.B., 2008. Errors in river lengths derived from raster digital elevation models. Comput. Geosci. 34 (11), 1584–1596. https://doi.org/10.1016/j.cageo.2007.10.009.

Poppenga, S.K., Worstell, B.B., 2016. Hydrologic connectivity: quantitative assessments of hydrologic-enforced drainage structures in an elevation model. J. Coast. Res. 76 (10076), 90–106. https://doi.org/10.2112/SI76-009.

Pravalika, P., Kumar, P.K., Srisaila, A., 2023. Bridge Detection using Satellite Images. 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) 1123–1129. https://doi.org/10.1109/ICAAIC56838.2023.10140305.

Qiu, M., Huang, L., Tang, B.-H., 2023. Bridge detection method for HSRRSIs based on YOLOv5 with a decoupled head. Int. J. Digit. Earth 16 (1), 113–129. https://doi.org/10.1080/17538947.2022.2163514.

Roelens, J., Rosier, I., Dondeyne, S., Van Orshoven, J., Diels, J., 2018. Extracting drainage networks and their connectivity using LiDAR data. Hydrol. Process. 32 (8), 1026–1037. https://doi.org/10.1002/hyp.11472.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A. F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.

Roussel, J.-R., Bourdon, J.-F., Morley, I.D., Coops, N.C., Achim, A., 2022. Correction, update, and enhancement of vectorial forestry road maps using ALS data, a pathfinder, and seven metrics. Int. J. Appl. Earth Obs. Geoinf. 114, 103020. https://doi.org/10.1016/j.jag.2022.103020.

Schneider, A., Jost, A., Coulon, C., Silvestre, M., Théry, S., Ducharne, A., 2017. Global-scale river network extraction based on high-resolution topography and constrained by lithology, climate, slope, and observed drainage density. Geophys. Res. Lett. 44 (6), 2773–2781. https://doi.org/10.1002/2016GL071844.

Schwanghart, W., Groom, G., Kuhn, N.J., Heckrath, G., 2013. Flow network derivation from a high resolution DEM in a low relief, agrarian landscape. Earth Surf. Process. Landf. 38 (13), 1576–1586.

Seibert, J., Grabs, T., Köhler, S., Laudon, H., Winterdahl, M., Bishop, K., 2009. Linking soil- and stream-water chemistry based on a Riparian Flow-Concentration Integration Model. Hydrol. Earth Syst. Sci. 13 (12), 2287–2297. https://doi.org/10.5194/hess-13-2287-2009.

Skoulikidis, N.T., Sabater, S., Datry, T., Morais, M.M., Buffagni, A., Dörflinger, G., Zogaris, S., del Mar Sánchez-Montoya, M., Bonada, N., Kalogianni, E., Rosado, J., Vardakas, L., De Girolamo, A.M., Tockner, K., 2017. Non-perennial Mediterranean rivers in Europe: status, pressures, and challenges for research and management. Sci. Total Environ. 577, 1–18. https://doi.org/10.1016/j.scitotenv.2016.10.147.

Soille, P., Vogt, J., Colombo, R., 2003. Carving and adaptive drainage enforcement of grid digital elevation models. Water Resour. Res. 39 (12). ://WOS:000187866800001.

Ståhl, G., Allard, A., Esseen, P.A., Glimskär, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Torell, Å.G., Högström, M., Lagerqvist, K., Marklund, L., Nilsson, B., Inghe, O., 2011. National Inventory of Landscapes in Sweden (NILS)-scope, design, and experiences from establishing a multiscale biodiversity monitoring system. Environ. Monit. Assess. 173 (1–4), 579–595. https://doi.org/10.1007/s10661-010-1406-7.

Stanislawski, L., Brockmeyer, T., Shavers, E., 2018. Automated road breaching to enhance extraction of natural drainage networks from elevation models through deep learning. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42 (4), 671–678. https://doi.org/10.5194/isprs-archives-XLII-4-597-2018.

Tiwari, T., Lidman, F., Laudon, H., Lidberg, W., Ågren, A.M., 2017. GIS-based prediction of stream chemistry using landscape composition, wet areas, and hydrological flow pathways. J. Geophys. Res.: Biogeosciences 122 (1). https://doi.org/10.1002/2016JG003399.

Van Nieuwenhuizen, N., Lindsay, J.B., DeVries, B., 2021. Automated mapping of transportation embankments in fine-resolution LiDAR DEMs. Remote Sens. 13 (7), 7. https://doi.org/10.3390/rs13071308.

Vaze, J., Teng, J., 2007. Impact of DEM Resolution on Topographic Indices and Hydrological Modelling Results. Modsim 2007: International Congress on Modelling and Simulation 706–712.

Wallin, B., Campeau, A., Audet, J., Bastviken, B., Bishop, K., Kokic, J., Laudon, H., Lundin, E., Löfgren, S., Natchimuthu, S., Sobek, S., Teutschbein, C., Weyhenmeyer, A.B., Grabs, T., 2018. Carbon dioxide and methane emissions of Swedish low-order streams—a national estimate and lessons learnt from more than a decade of observations. Limnology and Oceanography 3 (3). https://doi.org/10.1002/lol2.10061, 3.

Wang, C.-K., Fareed, N., 2021. Mapping drainage structures using airborne laser scanning by incorporating road centerline information. Remote Sens. 13 (3), 3. https://doi.org/10.3390/rs13030463.

Wang, L., Liu, H., 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. Int. J. Geogr. Inf. Sci. 20 (2), 193–213.

White, R.A., Dietterick, B.C., Mastin, T., Strohman, R., 2010. Forest roads mapped using LiDAR in steep forested terrain. Remote Sens. 2 (4), 4. https://doi.org/10.3390/rs2041120.

Wu, D., Li, R., Rekabdar, B., Talbert, C., Edidem, M., Wang, G., 2023. Classification of drainage crossings on high-resolution digital elevation models: a deep learning approach. GIScience Remote Sens. 60 (1), 2230706. https://doi.org/10.1080/15481603.2023.2230706.

Yang, W., Gao, X., Zhang, C., Tong, F., Chen, G., Xiao, Z., 2021. Bridge extraction algorithm based on deep learning and high-resolution satellite image. Sci. Program. 2021, e9961963. https://doi.org/10.1155/2021/9961963.