Environmental Modelling and Software 140 (2021) 104992

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: http://www.elsevier.com/locate/envsoft

An attention U-Net model for detection of fine-scale hydrologic streamlines

Zewei Xu [a,b], Shaowen Wang [a,b,*], Lawrence V. Stanislawski [c], Zhe Jiang [d], Nattapon Jaroenchai [a,b], Arpan Man Sainju [d], Ethan Shavers [c], E. Lynn Usery [c], Li Chen [b,e], Zhiyu Li [a,b], Bin Su [a,b]

[a] Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[b] CyberGIS Center for Advanced Digital and Spatial Studies, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[c] U.S. Geological Survey, Center of Excellence for Geospatial Information Science, Rolla, MO, USA
[d] Department of Computer Science, University of Alabama, Tuscaloosa, AL, USA
[e] School of Geosciences and Info-Physics, Central South University, Changsha, Hunan, China

ARTICLE INFO

Keywords:
CyberGIS
Deep learning
Hydrologic streamlines
Hydrography
Lidar data analysis

ABSTRACT

Surface water is an irreplaceable resource for human survival and environmental sustainability. Accurate, finely detailed cartographic representations of hydrologic streamlines are critically important in various scientific domains, such as assessing the quantity and quality of present and future water resources, modeling climate changes, evaluating agricultural suitability, mapping flood inundation, and monitoring environmental changes. Conventional approaches to detecting such streamlines cannot adequately incorporate information from the complex three-dimensional (3D) environment of streams and land surface features. Such information is vital to accurately delineate streamlines. In recent years, high accuracy lidar data has become increasingly available for deriving both 3D information and terrestrial surface reflectance. This study develops an attention U-net model to take advantage of high-accuracy lidar data for finely detailed streamline detection and evaluates model results against a baseline of multiple traditional machine learning methods. The evaluation shows that the attention U-net model outperforms the best baseline machine learning method by an average F1 score of 11.25% and achieves significantly better smoothness and connectivity between classified streamline channels. These findings suggest that our deep learning approach can harness high-accuracy lidar data for fine-scale hydrologic streamline detection, and in turn produce desirable benefits for many scientific domains.

1. Introduction

Interactions of water within Earth's systems have been studied extensively, yet increased demand for this vital resource has expanded interest in monitoring and management of water resources. Accurate, finely detailed delineation of surface hydrologic features is crucial for various scientific investigations and water resource applications, such as agricultural suitability, river dynamics, flood mapping, landslide risk analysis, wetland inventory, watershed analysis, environmental monitoring, and climate modeling, to name just a few (Maidment, 2017; Poppenga and Gesch, 2013; Schultz et al., 2017; Simley and Carswell, 2009; Terziotti, 2018; Wright and Nielsen, 2012). While other terrain conditions have a role, the spatial pattern of a surface water drainage network is largely a reflection of the type and arrangement of subsurface bedrock, which can assist with classification and management of land resources (Clubb and Bookhagen, 2019; Muller and Oberlander, 1976). Therefore, the key objective of this research is to understand how to advance machine intelligence for automatic extraction of detailed hydrologic features from high-resolution elevation data and other open geospatial datasets, yielding important data that can be used for this type of scientific work.

The National Hydrography Dataset (NHD) is a digital database of surface water features of the United States that is managed by the U.S. Geological Survey (USGS) and partner organizations (Sheng and Wilson, 2007; Simley and Carswell, 2009). It provides a common reference for regulation, research and modeling (NOAA, 2016). The NHD High Resolution (HR) is a multi-scale dataset compiled from the best available data sources having scales of 1:24,000 or larger (finer detail), except in Alaska where 1:63,360 or larger scales (finer details) are used. However, the quality of hydrographic data that has been compiled from

* Corresponding author. Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA.
E-mail address: shaowen@illinois.edu (S. Wang).

https://doi.org/10.1016/j.envsoft.2021.104992
Accepted 13 February 2021

topographic maps, which include the NHD, is not suitable for certain hydrologic, regulatory, and engineering purposes because of inconsistent drainage density and missing headwater content (Caruso, 2014; Chorley and Dale, 1972; Colson and Gregory, 2008; Colson, 2006; Fritz et al., 2013; Russell, 2008). Headwaters are small streams formed at the upstream extent of a watershed and comprise more than 50 percent of the stream network by length in the United States (Nadeau and Rains, 2007). To overcome these issues, since 2009 NHD HR is being updated with more detailed hydrography derived from finer-scale source information, up to 1:2400 (Simley and Carswell, 2009; Stanislawski, 2009). Economically, these enhanced hydrographic data are expected to generate over 600 million dollars per year in potential benefits to water resource and emergency response managers, in addition to the 500 million dollars in annual benefits already being generated from the existing program (Hoegberg, 2016).

Updating the NHD HR applies the best available digital elevation model (DEM) data, which should use Quality Level 2 (QL2) lidar data or better in the conterminous United States (Heidemann, 2018). Since 2014, the USGS 3D Elevation Program (3DEP) has been coordinating the collection of QL2 or better lidar point cloud data for the United States, except for Alaska where cloud-penetrating interferometric synthetic aperture radar (ifsar) is being acquired to simplify collection in remote areas (Lukas et al., 2015). QL2 lidar provides an aggregate nominal pulse spacing of less than 1-m (m) for first returns (Heidemann, 2018), which supports derivation of a 1 meter (m) resolution DEM. The detail inherent to this high-resolution DEM data enables modeling of surface water dynamics from the continental scale to catchment and headwater scales.

Although some methods to improve the NHD HR have been studied (Lopez-Torrijos, 2018; Poppenga et al., 2013; Sheng et al., 2007; Stanislawski and Survila, 2018), extracting accurate and fine-scale hydrography from high-resolution DEM data using traditional flow accumulation methods is a costly and laborious process. Depending on the selected workflow, various sophisticated issues must be handled, which include conditioning the DEM for flow modeling, estimating flow accumulation weights and a minimum contributing area for stream formation, along with tailoring solutions to diverse environmental conditions. Coupled with the fact that multiple methods are available, solutions can vary, and assessing the accuracy of extracted drainage lines is further complicated by temporal environmental variations. Procedures generally involve well-known automated methods to derive drainage lines from DEMs (Anderson, 2012; Jenson and Domingue, 1988; Maidment and Morehouse, 2002; Metz and Mitasova, 2011; Montgomery and Foufoula-Georgiou, 1993; O'Callaghan and Mark, 1984; Passalacqua and Belmont, 2012; Poppenga et al., 2013; Tarboton and Bras, 1991), with subsequent manual editing to adjust drainage lines and collect waterbodies from high-resolution orthorectified images.

Remotely sensed information at high spatial and temporal resolutions, such as repeat lidar, can facilitate automated analysis and extraction of hydrographic features, saving time and increasing the accuracy and consistency of extracted features (Sharma and Xu, 2016). Advanced computationally-intensive machine learning approaches integrated with cyberGIS (cyber geospatial information science and systems) for the resolution of computational and data intensive geospatial analyses (Wang, 2010; Wang and Goodchild, 2019; Wang and Liu, 2016), represent an exciting frontier for extracting accurate and fine-scale hydrography from lidar to improve the NHD HR.

Recent rapid advances in deep learning have been widely acknowledged and adopted in many challenging pattern recognition and object detection tasks (Kampffmeyer and Salberg, 2016; LeCun and Bengio, 2015; Maggiori and Tarabalka, 2017; Reichstein et al., 2019; Schmidhuber, 2015; Sun and Zhang, 2018; Xu and Guan, 2018; Zhu et al., 2017). Compared to the traditional or hand-crafted feature engineering, deep learning has demonstrated advances in accuracy and efficiency for complex feature learning in various application domains (Liang and Sun, 2017; Lin and Tegmark, 2017; Lin and Nie, 2017; Xu and Mountrakis, 2017). While such strategies promise a new way for hydrologic feature extraction from geospatial big data, limited effort has taken advantage of deep learning for accurate, efficient, and fine-scale delineation of hydrologic features. Moreover, the full utilization of the most recent technology of Geiger-mode lidar could significantly improve high-quality delineation of natural features (Clifton et al., 2015; Stoker and Abdullah, 2016).

This research develops a deep learning model based on the U-net structure (Ronneberger and Fischer, 2015) and attention mechanism (Oktay et al., 2018; Vaswani et al., 2017), which consists of a contractive path and an expanding path for segmenting streamlines from input feature maps. The contractive path is comprised of six triple convolutional layers plus five pooling layers for accurate extraction of global features and reduction of spatial redundancy, while the expanding path is comprised of five transposed convolutional layers plus five triple convolutional layers for projecting the extracted global feature content to original locations in the prediction map. The number of layers is chosen to reduce (contractive path) and upsample (expanding path) the x-y dimension of feature maps from 224-by-224 to 7-by-7 or from 7-by-7 to 224-by-224 with a stride of two. The patch size of 224 influences the model accuracy and efficiency. Small patch sizes cause poor accuracies due to the lack of context information while large patch sizes add extra computational burden without particular benefits to model performance. In this research, patch sizes of 64, 112, 224, 448, and 512 were tested and 224 was chosen over smaller or larger ones based on evaluation of model accuracy and efficiency. Meanwhile, feature concatenation is used to combine the extracted local and global information at different levels from the contractive path to its corresponding locations in the expanding path, to enhance the expressivity of the model during the convolution and transposed convolutional processes.

Computationally, we use GPU processing to speed up model training that is based on Keras and TensorFlow. Two types of benchmark methods are adopted for model comparisons. The first type includes two traditional pixel-based classification methods – a Support Vector Machine (SVM) and an Artificial Neural Network (ANN) model. The other types include the NHD HR data compiled from topographic maps and orthophotography, and elevation-derived drainage lines generated from GeoNet tools (Sangireddy and Stark, 2016). The comparison shows that our method based on the attention U-net model outperforms the best benchmark method by 8.61%, 9.39%, 13.68%, and 13.31% in four different scenarios. The resulting streamline map also indicates that the attention U-net model generates smoother and more topologically connected features than the benchmark methods, which is significant for hydrologic applications. The major contributions of this research are two-fold: (1) a novel application of the attention U-net for accurate and fine-scale hydrologic streamline detection, and (2) an effective streamline detection method that fully utilizes the geometric and intensity information from high-resolution lidar data.

2. Study area and input dataset development

Our study area is a watershed in Rowan County, which is located in west central North Carolina (Fig. 1). This area encompasses a set of tributaries that flow into Second Creek, which is the primary flowline feature of 12-digit NHD watershed 030401020504. The study area is 6.3 km^2 and has a humid, subtropical climate. Winters are short and mild, while summers are usually hot and humid. Spring and fall are distinct and refreshing periods of transition. Temperature ranges between 100°F (38°C) to 10°F (−12°C) within a year. The watershed lies in the Central Interior and Appalachian ecological division (Comer et al., 2003). In terms of land cover types, forest dominates the area and most of the stream channels are underneath closed-canopy. The lidar dataset used in this research is small-footprint, discrete-return, Geiger-mode lidar that was collected by the state government of North Carolina in the fall of 2016. The lidar dataset requires about 21 GB of disk storage, and the

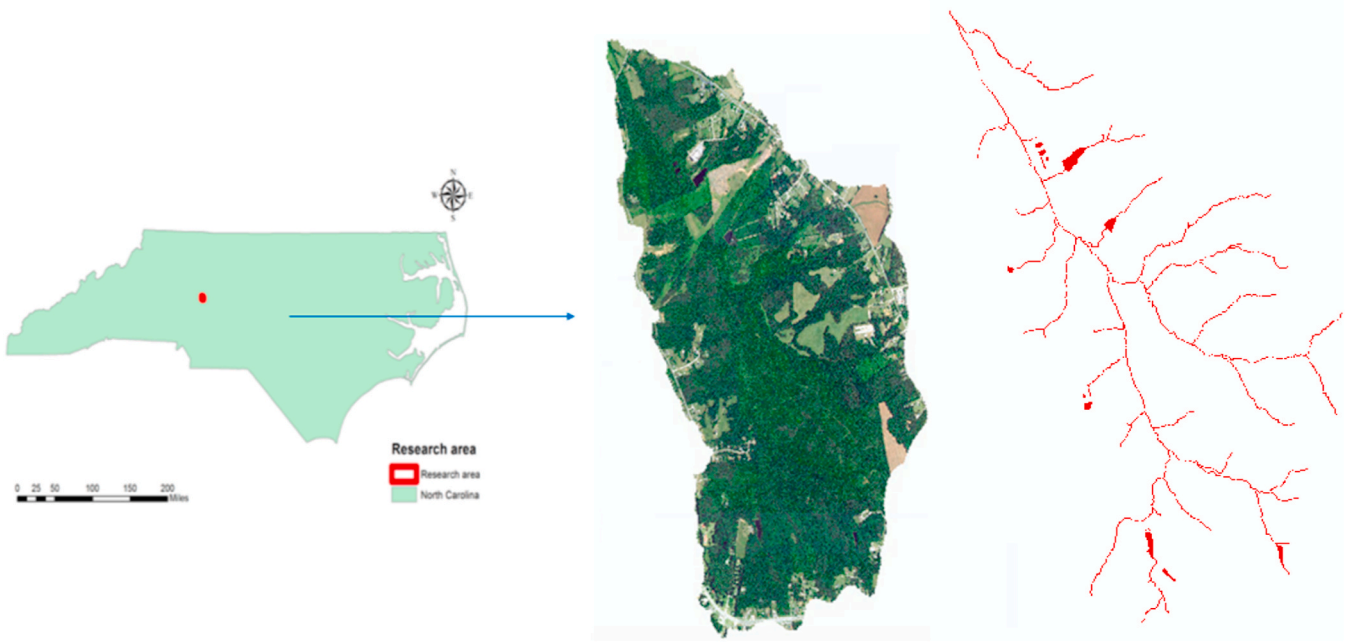

Fig. 1. Research area (Left: boundary of North Carolina state; middle: a 1-m resolution image from the National Agriculture Imagery Program; right: reference data).

projection coordinate system is the 2011 State Plane of North Carolina. This area has elevations ranging from 194 to 256 m. Because a Geiger-mode lidar sensor was used, the point density of all returns reaches 43 returns per square meter. A field validated set of intermittent stream heads surveyed between 2013 and 2014 along with on-screen editing was used to generate reference data (Shavers and Stanislawski, 2018). A 3-m buffer was generated along the reference streamline to simulate the width of stream channels.

Eight co-registered 1-m resolution raster data layers were derived from the lidar point cloud data and were used for training, validation, and testing in the research. The layers were selected through extensive comparison of elevation derivatives and optical imagery having national coverage with validated surface hydrography in diverse landscapes. The raster layers include: (1) a 1-m resolution digital elevation model (DEM) derived from the ground return points; (2) geometric curvature determined from the DEM; (3) a topographic position index (TPI) derived from the DEM using a 3-cell by 3-cell window; (4) a TPI derived from the DEM using a 21-cell by 21-cell window; (5) zenith angle positive openness derived from the DEM using a 10-cell radius with 32 directions (Doneus, 2013); (6) return intensity determined from the lidar ground points averaged with inverse distance weighting using 10 nearest points; (7) point density for return points between zero and 1 foot above ground; and (8) point density for return points between 0 and 3 feet above ground. Geometric curvature is determined using GeoNet software (Sangireddy et al., 2016). The software applies the non-linear diffusion Perona-Malik filter on the DEM to remove noise and sharpen the localization of channels (Passalacqua et al., 2010). Geometric curvature, which sums curvature in the x and y directions, is then determined for the filtered DEM. The TPI value of a cell is the difference between the cell elevation and the local average elevation within a specific radius or within a surrounding window of cells (De Reu et al., 2013). As noted, average values for TPI layer (3) and (4) are computed based on 3×3 and 21×21 surrounding cell windows, respectively. The TPI exaggerates local lows and highs in a DEM relative to the nearby topographic features, accentuating ridges and valleys. Zenith angle positive opennes (5) with 10-m radius can enhance drainage and small stream channels (Doneus, 2013). Lidar return intensity (6) is usually lower for water surfaces and wet areas than for dry areas because of energy absorption by water (Hooshyar et al., 2015). Return point density layers (7) and (8) respectively estimate the density of land surface features such as shrubs and tree limbs up to 1 and 3 feet above ground, which is most likely vegetation under the forest canopy. Shavers and Stanislawski (2018) suggest vegetation density structure in the riparian zones may be reflected in these layers. The eight raster layers are shown in Fig. 2 with summary statistics presented in Table 1.

Inputs to our model are individual image patches sampled from the eight different feature maps derived from the lidar data at a resolution of 1 m. The feature maps are normalized versions of the eight raster data layers, normalizing each of the floating point datasets within the study area to a corresponding unsigned integer feature map. To effectively test our method, we created four different classification scenarios by splitting the research area into upper/lower and left/right portions. When conducting our experiments, we used one of the portions for generating training/validation patches, and the other portion to generate testing patches for accuracy assessment in order to evaluate model generalizability. Sample patches for training and validation were generated based on a random process that ensures no overlap between training and validation patches. A visualization of the locations of our generated training and validation patches is shown in Fig. 3. To further enhance our training data, we applied image augmentation by randomly rotating each training patch by 30–150 and 210–330°, rescaling each sample by 0.5–0.8 and 1.5–2.0, shearing each sample by random ranges from −30 to 30°, and mirroring each sample horizontally to create six augmented samples for each training sample. Finally, 200 training (1400 after data augmentation) and 30 validation patches were selected.

3. Methods

3.1. Benchmark methods

To evaluate the performance of our method, it is compared to existing hydrography data, elevation-derived drainage lines, and hydrography predicted from two machine learning methods. We set up four baseline benchmarks, which include NHD HR, elevation-derived drainage lines from GeoNet, SVM, and an ANN model. The high resolution NHD HR was retrieved from the USGS NHD website (https://www.usgs.gov/core-science-systems/ngp/national-hydrography). NHD data for the study area were compiled from 1:24,000-scale digital

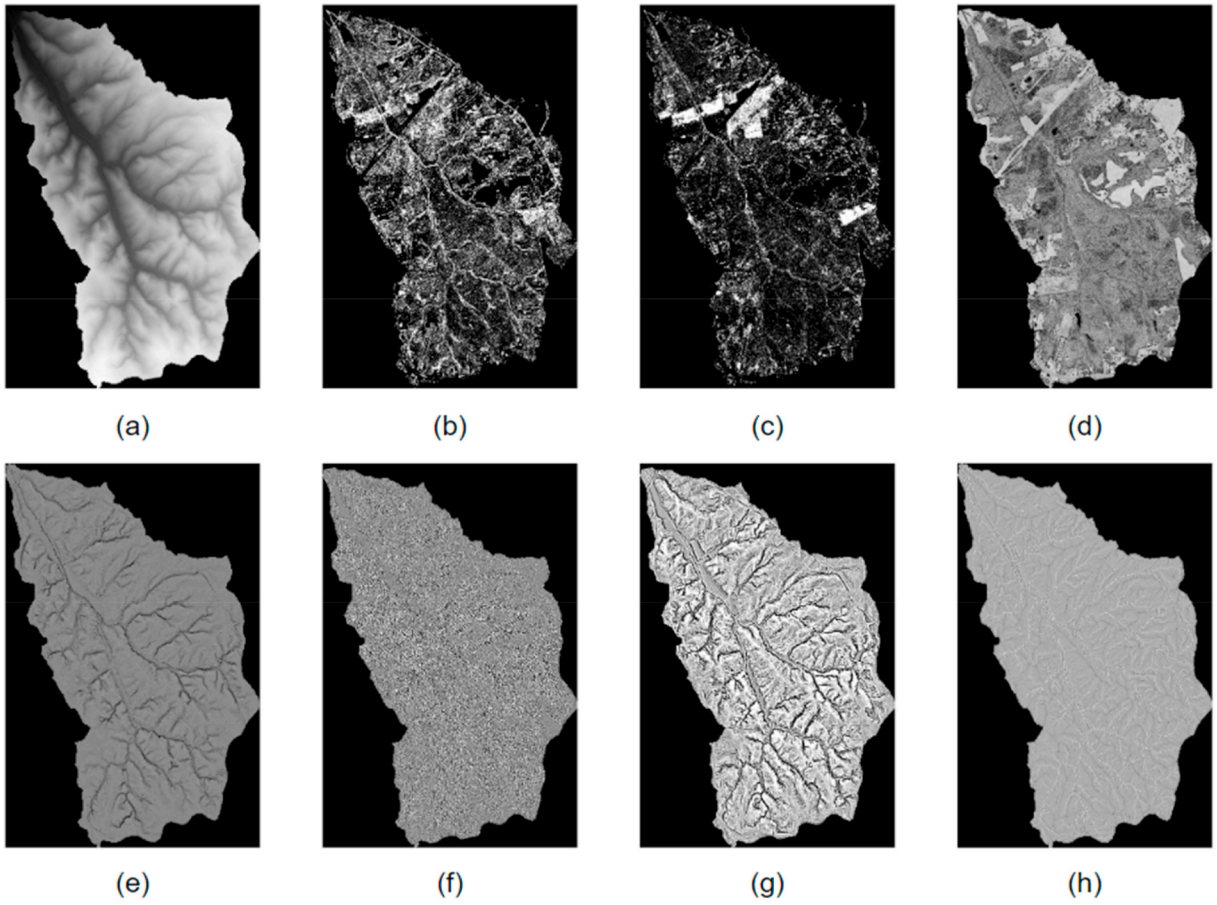

Fig. 2. Lidar feature maps: (a) Digital Elevation Model (DEM); (b) Digital Surface Model (DSM) of returns up to 1 feet above ground; (c) Digital Surface Model (DSM) of returns up to 3 feet above ground; (d) Lidar Reflectance; (e) Positive openness; (f) TPI with moving window size 3; (g) TPI with moving window size 21; (h) Topological Curvature.

line graph data in 2001, with waterbodies and associated features manually adjusted in 2013 to fit National Agriculture Imagery Program (NAIP) 1-m resolution color-infrared digital orthophotography. The GeoNet lines are extracted from the DEM using a least-cost path-tracing technique that is guided by a minimum threshold flow accumulation skeleton (Sangireddy et al., 2016). In our case, the flow accumulation skeleton is generated using a minimum threshold of 1000 cells, which is expected to over extract water flow network and fully define the drainage paths.

We used a SVM classifier based on a radial basis function (RBF) kernel with a kernel approximation strategy for speeding up the training process (Rahimi and Recht, 2008). The parameters of kernel degree (g) and penalty (C) are tuned using a two-level grid search in the range of 10-5 to 105 and 10-5 to 1, respectively. For the benchmark neural network model, we construct a model with two hidden layers and a sigmoid activation function as the output layer. The parameters of number of hidden layers, learning rate, momentum, and decaying rate are also tuned using grid search. The reference data including training, validation, and testing data are the same between different models for model training, parameter tuning and generating the final feature maps.

3.2. The U-net model

The U-net model is a special type of Fully Convolutional Networks (FCNs). Unlike normal convolutional neural networks (CNNs), the last fully connected layer from FCNs are substituted by a series of transposed convolutional layers with larger and larger receptive fields. FCNs are built only by locally connected layers including convolution, pooling and upsampling layers without using any dense connected layer. This practice greatly reduces the number of parameters for model tuning and thus reduces redundant computation compared to traditional CNNs. A typical FCN has two parts: a contractive path and an upsampling path, where the former is used to extract important information and reduce spatial redundancy, and the latter is used to project the extracted information to specific locations in the original image (Ronneberger et al., 2015). The U-net model is a state-of-the-art FCN that achieves high accuracy for solving image segmentation problems (Ronneberger et al., 2015). Based on the fundamental structure of an FCN, it further applies feature concatenations to recover and fully utilize the information extracted at different resolution levels in the contractive path to the corresponding locations in the expanding path. Details of model layers are described as follows.

Convolutional layer: the convolutional layer is the major workforce for extracting important features from images (Krizhevsky and Sutskever, 2012). It conducts image filtering by using kernel filters. In this process, image features with strong signals are extracted.

Pooling layer: Pooling layer is used to conduct downsampling on activation maps. Downsampling reduces the sampling rate of a raster by decreasing the raster resolution (i.e., increasing pixel size). The max function is often used to filter out redundant information and preserve the strongest feature signals.

Relu layer: Relu layer is short for rectified linear unit layer, which is one of the most commonly used activation functions in CNNs (Agarap, 2018). It consists of a linear function for all positive input values, and zero for all negative values. It truncates unimportant features generated from the convolutional layer and only reserves the important ones.

Transposed convolutional layer: This layer projects the extracted dense features from the coarse resolution to its precise location in the

Table 1
Summary statistics for 1-m resolution raster datasets derived from lidar point cloud data collected for Rowan County watershed.

Raster Dataset	Minimum	Maximum	Mean	Standard Deviation	Range
Digital elevation model (meters)	194.11	256.19	229.07	12.96	62.07
Geometric curvature	−97.25	97.93	0.01	3.05	195.18
Topographic position index (3 × 3 window)	−8.59	5.58	6.38	0.18	14.167
Topographic position index (21 × 21 window)	−13.62	13.29	0	0.93	26.91
Openness (R10, D32) degrees	21.52	118.8	83.41	7.35	97.28
Return intensity	0	55185.39	29047.18	10624.11	55185.39
Return point density 1 ft above ground (points per m^2)	0	0.94	0.02	0.04	0.94
Return point density 3 ft above ground (points per m^2)	0	2.89	0.12	0.23	2.89

original image by using upsampling (i.e., increasing pixel resolution) or spatial interpolation.

The attention module is a technique originally designed for sequence dependency modeling that has recently been adopted for modeling feature dependencies in image analysis (Oktay et al., 2018; Vaswani et al., 2017). It can progressively suppress feature responses in irrelevant background regions and make the model focus on important features. In this research, we integrate five attention gates (AGs) into the U-net model and thus create an attention U-net model for achieving high accuracy results. As shown in Fig. 4a, the standard attention module maps query pixels and their key-value pairs to the output. The output is a set of weighted values and the attention weight matrix is calculated by a compatibility function of the query with the corresponding key (Vaswani et al., 2017). Finally, the weighted inputs are multiplied by a scaling hyperparameter α (initialized as 1) and added to the original input to produce the final output. Since the original attention weight matrixes at shallower layers are too large to fit in the memory, we used a second attention module (Fig. 4b) adapted from (Oktay et al., 2018) in our model. The only difference is the second one directly combines the convolutional results from the feature maps and the gating signal to a Relu layer to remove negative values, and utilizes a bottleneck convolutional layer to reduce the channel dimension for memory saving and a sigmoid function to calculate the final attention weight matrix.

The architecture of the attention U-net model is shown in Fig. 5. It applies six triple convolutional layers to the contractive path and five in the expanding path. Five pooling layers are used between each of the triple convolutional layers for downsampling. In the expanding path, five transposed convolutional layers (size 2 × 2 and stride 2) are used for feature upsampling. In each horizontal level of the two paths, the network uses attention gates to filter the features propagated through the skip connections based on the gating signal of the contextual information from coarser scales to achieve high accuracy in segmentation results. We utilize an "Adam" optimizer for calculating the change direction of loss and adjust the weights in the back-propagation process (Kingma and Ba, 2014).

We use Python 2.7 and Keras 2.0 with backend of TensorFlow 1.0 for the model construction. We also utilize Python libraries including sklearn 0.18.1, scikit-image 0.16.2, GDAL 3.0.2, NumPy 1.17.3 supported by Anaconda 2.0. The model is tested using both GPU and CPU devices. It takes 15 h using a state-of-the-art CPU and 2 h for a Tesla M80 GPU to finish model training for our 6.07 km^2 study area. In this paper, we run the attention U-net and the U-net model separately five times and the average statistics are reported for evaluation.

The attention U-net model uses Dice's coefficient (1) as the loss function. The coefficient is the quotient of similarity and ranges between 0 and 1. Dice's coefficient value equals twice true positive (TP) divided by the sum of twice true positive (TP), false positive (FP), and false negative (FN) as shown in equation (1). We use precision, recall, and F1 score to evaluate the model performance against testing data. Because of the difficulty in correctly labeling all streamline pixels, relaxed methods are adopted to calculate precisions and recalls (Mnih and Hinton, 2010). The relaxed precision is defined as the fraction of number of pixels predicted as stream within a range of ρ pixels from pixels labeled as stream. The relaxed recall is the fraction of number of pixels labeled as stream that are within a range of ρ pixels from pixels predicted as stream pixels. In our experiments, the slack parameter ρ is set to 3 according to previous research (Mnih and Hinton, 2010).

$$Dice's\ coefficient = \frac{2\mathrm{TP}}{2TP + FP + FN} \tag{1}$$

We used a grid search for hyperparameter tuning of the learning rate, filter size, dropout rate, and decaying factor. In this process, the "Adam" optimizer is used to calculate and adjust the weights during training (Zhang, 2018). Fig. 6 shows the change of training accuracy from different learning rates against the number of training epochs using the standard U-net model. We can see that the learning rate of 3.59e-05

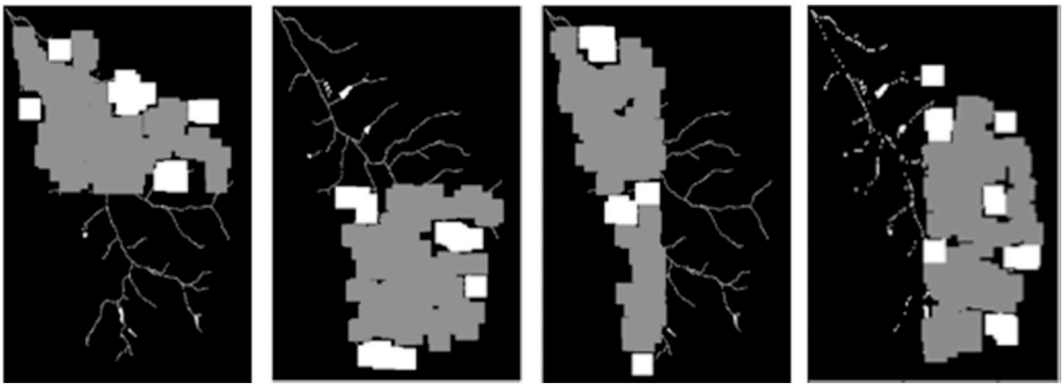

Fig. 3. Distribution of training and validation patches in the four scenarios (gray area: locations of training patches; white area: locations of validation patches).

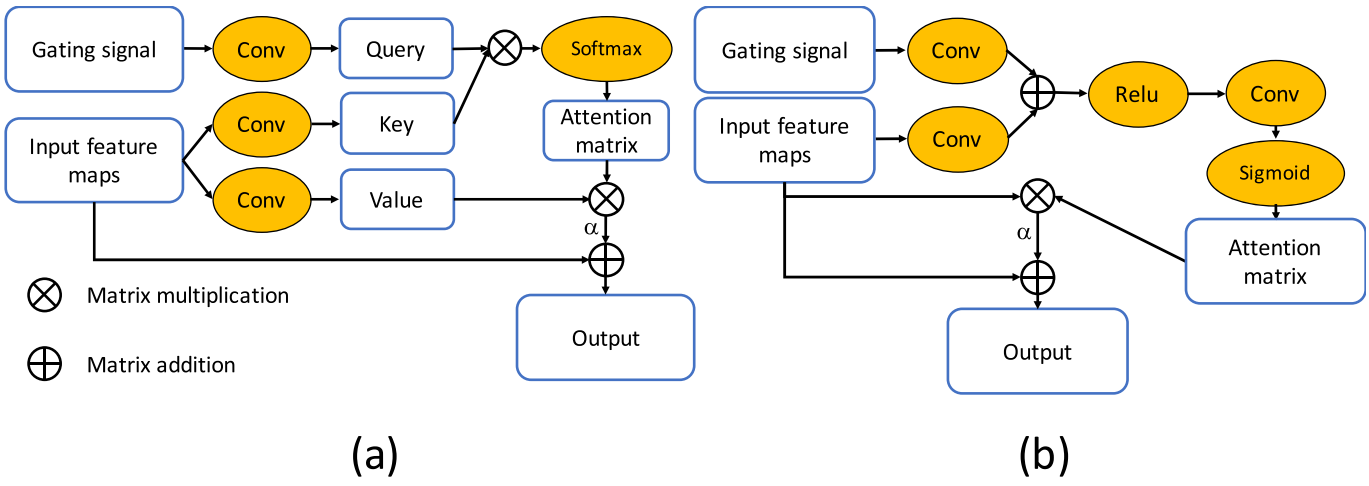

Fig. 4. Attention gates of the U-net model (a: attention gate 1; b: attention gate 2).

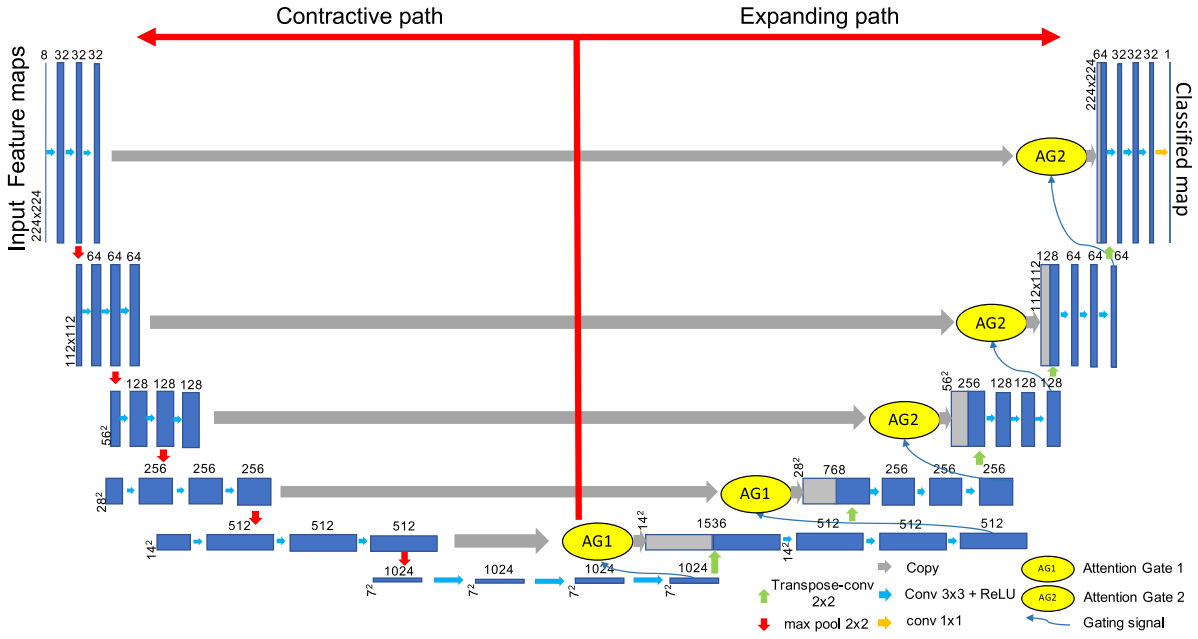

Fig. 5. The architecture of the attention U-net model.

achieves a good convergence and accuracy. The plot of training and validation losses using the selected learning rate is shown in Fig. 7.

4. Results

We evaluate our method against multiple benchmark methods in four scenarios. For the first two scenarios, we split the study area horizontally and use the lower portion as testing data in scenario one, and the upper portion as testing data in scenario two. For scenarios three and four, we split the study area vertically and use the right portion as testing data in scenario three, and the left portion as testing data in scenario four. Three metrics, precision (2), recall (3), and F1 score (4) are used to evaluate the performance of the methods and are defined as follows (TP: True Positive; FP: False Positive; FN: False Negative).

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1\ score = 2^{*}\frac{Precision\ *\ Recall}{Precision\ +\ Recall} \tag{4}$$

In this research, we focus on the F1 score because it is the harmonic mean of precision and recall. The highest F1 score means the model has an optimal balance of recall and precision. While recall expresses the model's ability to find all streamline pixels in the input data, precision expresses the portion of pixels that a model classifies as streamline correctly. Therefore, there is a trade-off between these two metrics.

The F1 score, precision, and recall of test accuracies for the attention U-net, U-net, and benchmark methods among the four scenarios are shown in Tables 2–4 respectively. Table 2 shows that both the attention U-net and U-net models outperform all benchmark methods for the four scenarios in overall F1 score. Also, the attention U-net slightly outperforms the U-net model in terms of the average F1 score. Among the

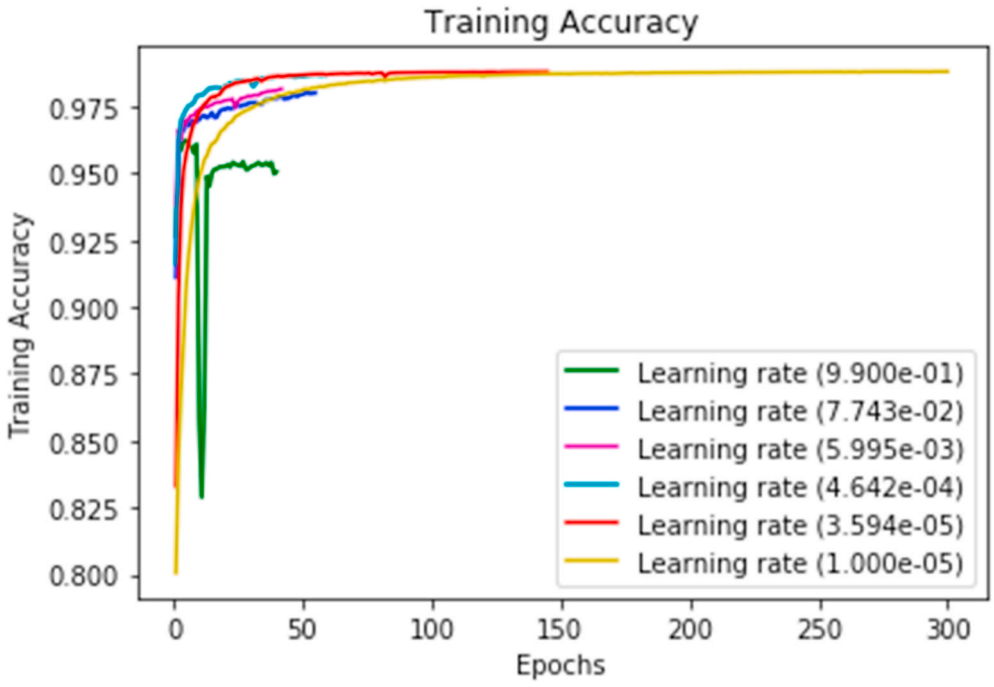

Fig. 6. Training accuracies with different learning rates.

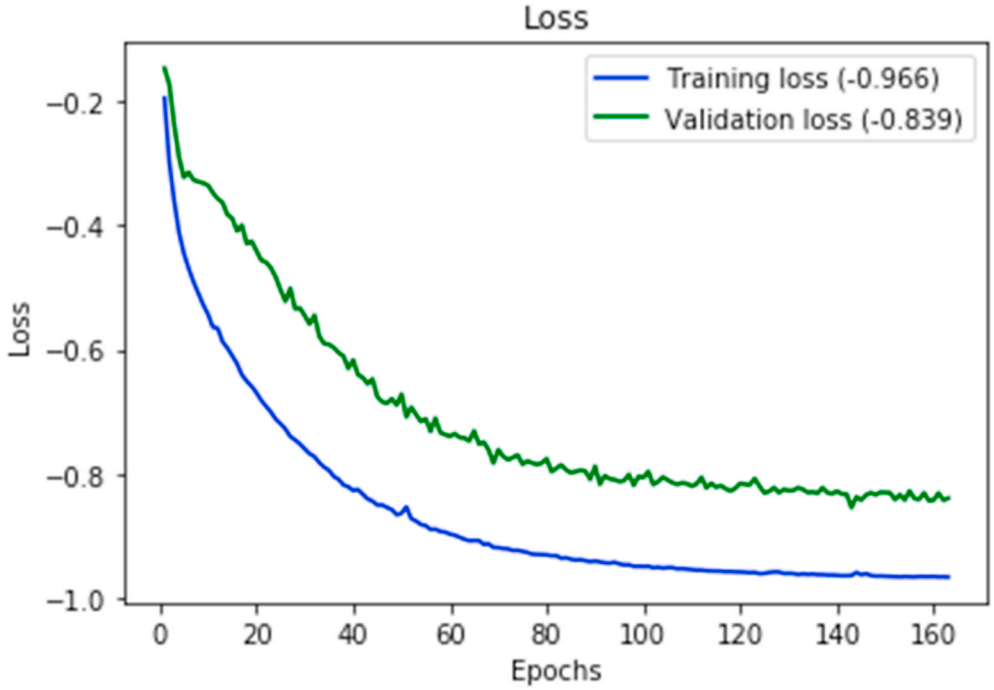

Fig. 7. Loss of the training and validation against epochs.

four benchmark methods, ANN achieves the highest accuracy and NHD has the lowest accuracy. The attention U-net model outperforms ANN by 8.61%, 9.39%, 13.68%, and 13.31%; SVM by 12.12%, 12.98%, 22.69%, and 13.54%; NHD by 45.51% and GeoNet by 23.61% on average. For precision, SVM achieves the best and outperforms the attention U-net model by 12.04%, 7.64%, 1.99%, and 9.51% from scenario 1 to 4. The attention U-net model outperforms the U-net model by 4.73%, 3.23%, 3.18%, and 2.84% from scenario 1 to 4. For recall, the GeoNet model has the highest accuracy of 92.66%, which is 1.78% higher than the attention U-net model. Apart from that, the attention U-net model achieves the next highest recalls and outperforms the U-net model by 0.64%, SVM by 32.15%, ANN by 17.62%, and NHD by 54.92% on average in scenario 1 to 4. Overall, the attention U-net model outperforms the U-net and all benchmark methods according to the average F1-score. Although one of the benchmark methods generates better precision values, errors reflected by recall values are large and makes it worse in terms of general performance compared to both the attention U-net and U-net model.

We visualize two large-extent locations (Fig. 8) and two further zoomed-in contexts (Fig. 9) in scenario 1 to demonstrate the improved performance compared with the benchmark methods. Both the attention

Table 2
Comparison of the F1 scores between the attention U-net, the U-net and benchmarks.

Scenarios (F1 score)	Attention U-net model	U-net model	SVM	ANN	NHD	Geonet
Scenario 1	83.02%	78.94%	70.90%	74.41%	41.05%	62.95%
Scenario 2	90.53%	87.94%	77.55%	81.14%		
Scenario 3	91.91%	90.61%	69.22%	78.23%		
Scenario 4	80.79%	79.82%	67.25%	67.48%		
Average	86.56%	84.33%	71.23%	75.32%		

Table 3
Comparison of the precisions between the attention U-net, the U-net and benchmarks.

Scenarios (Precision)	Attention U-net model	U-net model	SVM	ANN	NHD	Geonet
Scenario 1	74.53%	69.80%	86.57%	80.90%	47.84%	47.67%
Scenario 2	87.96%	84.73%	95.60%	82.40%		
Scenario 3	91.06%	87.88%	93.05%	82.34%		
Scenario 4	78.16%	75.32%	87.67%	65.38%		
Average	82.92%	79.43%	90.72%	77.76%		

Table 4
Comparison of the recalls between the attention U-net, the U-net and benchmarks.

Scenarios (Recall)	Attention U-net model	U-net model	SVM	ANN	NHD	Geonet
Scenario 1	93.80%	91.10%	60.04%	68.88%	35.96%	92.66%
Scenario 2	93.27%	91.41%	65.24%	79.92%		
Scenario 3	92.78%	93.53%	55.10%	74.50%		
Scenario 4	83.66%	84.93%	54.55%	69.73%		
Average	90.88%	90.24%	58.73%	73.26%		

U-net and U-net model generate better streamline delineations with better connectivity and smoother shapes following channels compared to ANN and SVM, which generate fragmented channels. The NHD vector features are smooth and well-connected, as are the GeoNet drainage lines, which are generated from a least-cost path model guided by flow accumulation (Sangireddy et al., 2016). However, GeoNet lines overestimate channels by false recognition of the dry drainage lines as stream channels. As expected from 1:24,000-scale data the NHD is sparse and only contains the several major channels in the study area. The U-net and attention U-net model also perform better in extracting most water-related features including water bodies in the two locations, where all the automated benchmark methods fail to do so. NHD includes the small lakes interpreted from orthophotography. When the attention U-net model is compared to the U-net model, the former eliminates many overestimated streamlines in the middle part of location 1 and 2, and better extracts water bodies. From Fig. 9, we can see that the attention U-net model is superior to traditional machine learning methods in extracting smooth streamlines and water bodies and avoids a majority of overestimated streamline pixels in NHD HR and the GeoNet flow accumulation model. The U-net model performs similarly to the attention U-net model, but the latter has a better delineation of the streamlines in the north and middle parts (less overestimations) of location 1 and better delineation of water bodies of location 2.

5. Discussion and conclusions

This research developed an attention U-net model for hydrologic streamline extraction using lidar-derived feature maps. Specifically, we have solved an image segmentation problem (segmentation of streamlines) based on the binary classification of stream versus non-stream pixels. This problem is difficult because hydrologic streamlines are formed by complex processes and occupy only a small portion of diverse land cover types while extracted streamlines need to be well connected. Furthermore, surface water features (e.g., clear/turbid rivers, swamps, ponds, and lakes) are spatially heterogeneous, and thus are difficult to extract using traditional machine learning methods (e.g., ANN and SVM) that cannot effectively handle multi-scale context information (e.g., topology, land cover distribution, topography).

The U-net model is a special type of fully CNNs using skip connections to combine local content from the contractive path to global content in the expanding path, which ensures adequate connectivity of segmentation results because it enables the model to take both the global and local context information into consideration while extracting streamlines. The attention module is added to the U-net model to

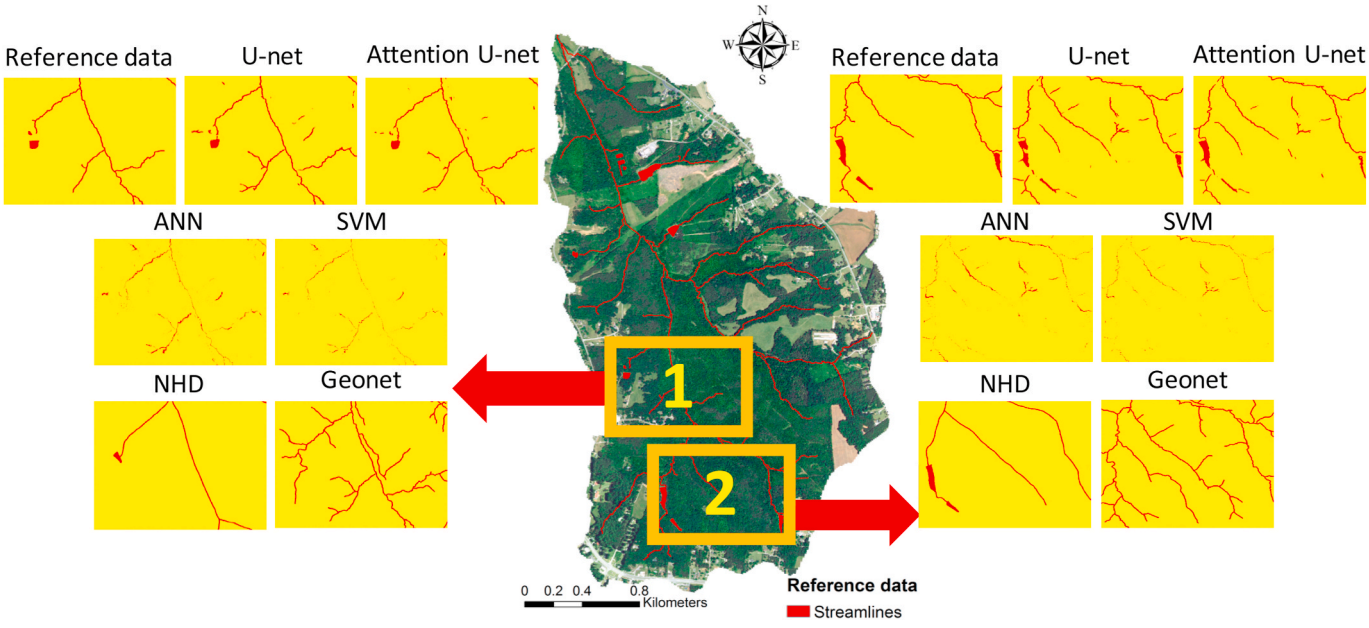

Fig. 8. Comparison of classification results from the attention U-net, U-net, and different benchmark methods.

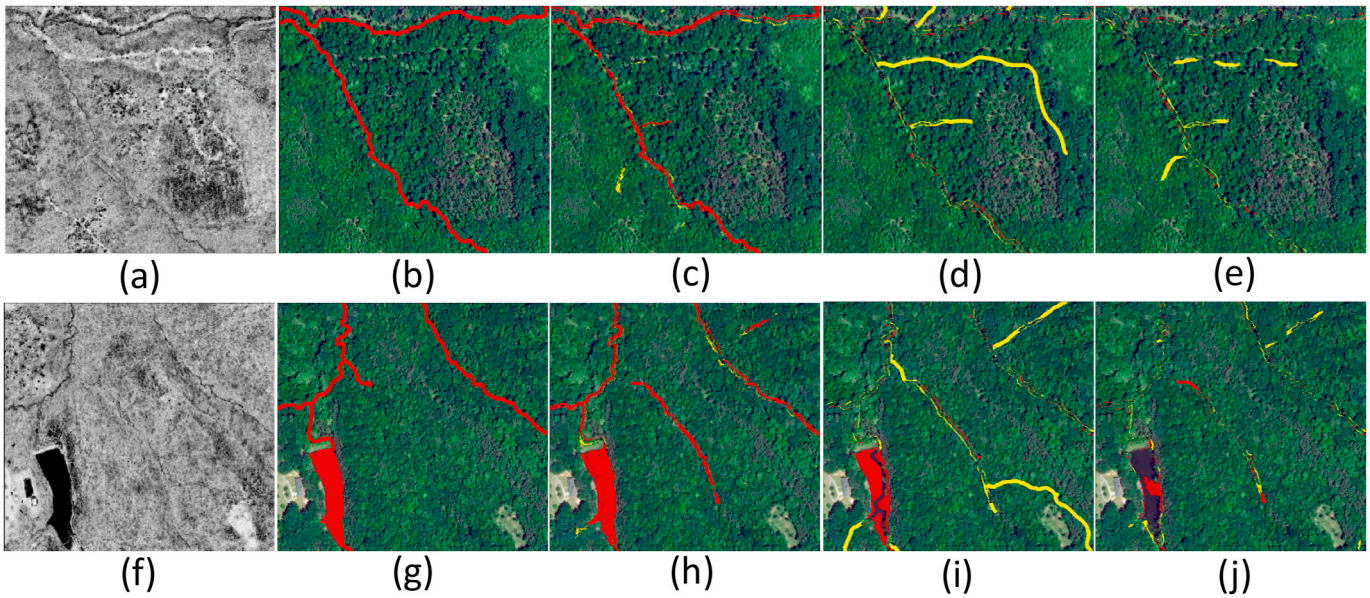

Fig. 9. Visualization of the feature differences between the attention U-net model versus the GeoNet and ANN models for two zoomed-in contexts: (a) and (f): Ground reflectance maps from lidar; (b) and (g): the reference data; (c) and (h): the difference map of the attention U-net model and the ANN model; (d) and (i): the difference map of the attention U-net model and the GeoNet model; (e) and (j): the difference map of the attention U-net model and the U-net model. Red areas are the pixels labeled as the stream feature by the attention U-net model and recognized as the non-stream feature by the benchmark models or the U-net model. The yellow areas are the pixels labeled as stream feature by the benchmark models or the U-net model and recognized as non-stream feature by the attention U-net model.

progressively suppress feature responses in irrelevant background regions and make the model focus on important streamline features according to the reference data. Addition of the attention module further enhances the accuracy for difficult instances such as the boundary of lakes, river bends, and dried channels.

A comprehensive evaluation of the model shows that our method outperforms multiple machine learning models and conventional flow accumulation methods by providing smoother and better-connected streamline and waterbody features. To evaluate the model thoroughly, we created four different scenarios by splitting our research area into upper/lower and left/right portions for generating training/validation and testing patches respectively. The attention U-net model generates F1-scores of 83.02%, 90.53%, 91.91%, and 80.79% across the four scenarios, which outperforms the best benchmark by 8.61%, 9.39%, 13.68%, and 13.31%.

Streamlines extracted using ANN and SVM are fragmented with missing parts mainly because these pixel-based classification methods fail to consider the global context. SVM achieves a high-precision result but a poor recall. This indicates that it underestimates the stream class pixels but provides good confidence of those extracted. Since the reference data are highly imbalanced (1:100 between stream and non-stream), we also conducted additional experiments that artificially upsample the stream samples. Although the imbalanced issue is resolved, the model has heavy overestimation of the stream class and the accuracy is not comparable to the imbalanced case. Compared to the manually verified reference data, the elevation-derived GeoNet drainage lines and features furnish higher recall scores than precision scores, which indicates that this method overestimates the stream class in general and performs better in terms of completeness than precision. These overestimations indicate local climate is drier than was assumed when selecting the flow accumulation threshold for the GeoNet lines, which include more dry tributaries than collected in the reference data. The 1:24,000-scale NHD only contains major stream channels and lake features and ignores the smaller tributaries, so the accuracy is much lower than the other datasets. Only the NHD benchmark data includes the water bodies in the study area, and the attention U-net model performs much better in terms of water body extraction than the other benchmark methods.

The attention U-net model utilizes its special feature concatenation design and a CNN to achieve high accuracy, adequate connectivity, and efficient streamline detection. Conventional machine learning models produce less optimal results primarily because they employ a pixel-based classification strategy. The attention U-net model also departs from traditional flow accumulation models that heavily rely on expert inputs, which in this case includes over-extracted drainage lines and no water body extraction leading to large errors. This research utilizes Geiger-mode lidar, which provides high-density point clouds and precise measurements for enabling transformative discovery and innovative opportunities in many scientific domains (McManamon et al., 2017).

The following set of principles distilled from this research are important for guiding the application of the method to other areas of study or solving similar problems.

- Ensure a balanced number of convolutional layers at each horizontal level in both the contractive and expanding paths. We find triple convolutional layers achieved adequate results and adding more convolutional layers would not benefit the model but increase computational intensity.
- Training patches should be randomly generated and have no overlap with validation and testing data. Training patches can overlap with themselves (effect of data augmentation).
- Use data augmentation by randomly rotating, mirroring, shearing, and rescaling training samples to ensure the expressivity of the model.
- A dropout layer and a proper dropping rate (hyperparameter) are necessary in the final convolutional process for the model regularization.
- Use early stopping to enhance the training efficiency and prevent overfitting of the model.

Future work will focus on applying the method to more study sites, and scaling the model up to regional and national scopes.

Disclaimer

Any use of trade, firm, or product names is for descriptive purposes

only and does not imply endorsement by the U.S. Government.

Author contributions

Z.X. and S.W. conceptualized the idea. S.W., L.V.S. and E.L.U. initialized, conceived and supervised the research. L.V.S. and E.S. collected data. Z.X., Z.J., A.M.S. and S.W. developed the deep learning model. Z.X., N.J., L.C., Z.L. and B.S. conducted experiments. Z.W., S.W., L.V.S., Z.J. and E.S. drafted the manuscript. All authors provided critical review of the manuscript and approved the final draft for submission.

Data availability

The data of this research is available for downloading from this link: https://doi.org/10.6084/m9.figshare.12584975.v1.

Code availability

The code is available under the open-source MIT license at: https://github.com/cybergis/streamline_detection.git.

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.

Acknowledgements

This paper and associated materials are based in part upon work supported by the National Science Foundation (NSF) under grant numbers: 1443080, 1743184, and 1833225. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF. Our computational work used Virtual ROGER, which is a cyberGIS supercomputer supported by the CyberGIS center for Advanced Digital and Spatial Studies and the School of Earth, Society and Environment at the University of Illinois at Urbana-Champaign.

References

Agarap, A.F., 2018. Deep Learning Using Rectified Linear Units (Relu) arXiv preprint arXiv:1803.08375.

Anderson, D.L., 2012. Detailed Hydrographic Feature Extraction from High-Resolution LiDAR Data. Doctoral Dissertation. Idaho State University), Idaho National Laboratory, U.S. Department of Energy National Laboratory. Retrieved June 19, from. https://inldigitallibrary.inl.gov/sites/sti/sti/5553757.pdf.

Caruso, B.S., 2014. GIS-Based stream classification in a mountain watershed for jurisdictional evaluation. JAWRA Journal of the American Water Resources Association 50 (5), 1304–1324. https://doi.org/10.1111/jawr.12189.

Chorley, R.J., Dale, P.F., 1972. Cartographic problems in stream channel delineation. Cartography 7 (4), 150–162. https://doi.org/10.1080/00690805.1972.10437698.

Clifton, W.E., Steele, B., Nelson, G., Truscott, A., Itzler, M., Entwistle, M., 2015. Medium altitude airborne Geiger-mode mapping LIDAR system. In: Laser Radar Technology and Applications XX; and Atmospheric Propagation XII, vol. 9465. International Society for Optics and Photonics, p. 946506.

Clubb, F.J., Bookhagen, B., Rheinwalt, A., 2019. Clustering river profiles to classify geomorphic domains. J. Geophys. Res.: Earth Surface 124, 1417–1439. https://doi.org/10.1029/2019JF005025.

Colson, T.P., 2006. Stream Network Delineation from High-Resolution Digital Elevation Models. PhD diss. North Carolina State University.

Colson, T., Gregory, J., Dorney, J., Russell, P., 2008. Topographic and soil maps do not accurately depict headwater stream networks. Natl. Wetl. Newsl. 30 (3), 25–28.

Comer, P., Faber-Langendoen, D., Evans, R., Gawler, S., Josse, C., Kittel, G., Menard, S., Pyne, M., Reid, M., Schulz, K., Snow, K., Teague, J., 2003. Ecological Systems of the United States, A Working Classification of U.S. Terrestrial Systems. NatureServe, Arlington, Va, p. 75.

De Reu, J., Bourgeios, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., Crombé, P., 2013. Application of the topographic position index to heterogeneous landscapes. Geomorphology 186, 39–49. https://doi.org/10.1016/j.geomorph.2012.12.015.

Doneus, M., 2013. Openness as visualization technique for interpretative mapping of airborne lidar derived digital terrain models. Rem. Sens. 5 (12), 6427–6442.

Fritz, K.M., Hagenbuch, E., D'Amico, E., Reif, M., Wigington Jr., P.J., Leibowitz, S.G., Comeleo, R.L., Ebersole, J.L., Nadeau, T.L., 2013. Comparing the extent and permanence of headwater streams from two field surveys to values from hydrographic databases and maps. JAWRA Journal of the American Water Resources Association 49 (4), 867–882. https://doi.org/10.1111/jawr.2013.49 issue-4.

Heidemann, H.K., 2018. Lidar Base Specification (Ver. 1.3, February 2018): U.S. Geological Survey Techniques and Methods, Book 11, p. 101. https://doi.org/10.3133/tm11b4 chap. B4.

Hoegberg, S., 2016. National Hydrography Requirements and Benefits Study - Preliminary Results. Dewberry LLC, Fairfax, Virginia, p. 140. https://www.usgs.gov/core-science-systems/ngp/national-hydrography/hydrography-requirements-and-benefits-study.

Hooshyar, M., Kim, S., Wang, D., Medeiros, S.C., 2015. Wet channel network extraction by integrating LiDAR intensity and elevation data. Water Resources Research 51, 10029–10046. https://doi.org/10.1002/2015WR018021.

Jenson, S.K., Dominigue, J.O., 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. Photogramm. Eng. Rem. Sens. 54 (11), 1593–1600.

Kampffmeyer, M., Salberg, A.B., Jenssen, R., 2016. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1–9.

Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization arXiv preprint arXiv:1412.6980.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436.

Liang, H., Sun, X., Sun, Y., Gao, Y., 2017. Text feature extraction based on deep learning: a review. EURASIP J. Wirel. Commun. Netw. 2017 (1), 1–12.

Lin, H.W., Tegmark, M., Rolnick, D., 2017a. Why does deep and cheap learning work so well? J. Stat. Phys. 168 (6), 1223–1247.

Lin, Y.Z., Nie, Z.H., Ma, H.W., 2017b. Structural damage detection with automatic feature-extraction through deep learning. Comput. Aided Civ. Infrastruct. Eng. 32 (12), 1025–1046.

Lopez-Torrijos, R., 2018. Best Practices in Hydrography Extraction from Lidar: Status Update. HydroShare. http://www.hydroshare.org/resource/7ec2a044ccf24482b1527858df669e19.

Lukas, V., Eldridge, D.F., Jason, A.L., Saghy, D.L., Steigerwald, P.R., Stoker, J.M., Sugarbaker, L.J., Thunen, D.R., 2015. Status Report for the 3D Elevation Program, 2013–2014: U.S. Geological Survey Open-File Report 2015–1161, p. 17. https://doi.org/10.3133/ofr20151161.

Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P., 2017. High-resolution aerial image labeling with convolutional neural networks. IEEE Trans. Geosci. Rem. Sens. 55 (12), 7092–7103.

Maidment, D.R., 2016. Conceptual framework for the national flood interoperability experiment. J. Am. Water Resour. Assoc. 53 (2), 245–257. https://doi.org/10.1111/1752-1688.12474.

Maidment, D.R. (Ed.), 2002. Archydro: GIS for Water Resources. ESRI Press, Redlands, CA.

Mcmanamon, P.F., Banks, P., Beck, J., Fried, D.G., Huntington, A.S., Watson, E.A., 2017. Comparison of flash lidar detector options. Opt. Eng. 56 (3), 031223 https://doi.org/10.1117/1.oe.56.3.031223.

Metz, M., Mitasova, H., Harmon, R., 2011. Efficient extraction of drainage networks from massive, radar-based elevation models with least cost path search. Hydrol. Earth Syst. Sci. 15, 667–678.

Mnih, V., Hinton, G.E., 2010. Learning to detect roads in high-resolution aerial images. In: European Conference on Computer Vision. Springer, Berlin, Heidelberg, pp. 210–223.

Montgomery, D.R., Foufoula-Georgiou, E., 1993. Channel network source representation using digital elevation models. Water Resour. Res. 29 (12), 3925–3934. https://doi.org/10.1029/93WR02463.

Muller, R.A., Oberlander, T.M., 1978. Physical Geography Today: Portrait of a Planet. Random House, New York, New York, p. 590pp.

Nadeau, T., Rains, M.C., 2007. Hydrological connectivity between headwater streams and downstream waters: how science can inform policy. J. Am. Water Resour. Assoc. 43 (1), 118–133. https://doi.org/10.1111/j.1752-1688.2007.00010.x.

NOAA (National Oceanic and Atmospheric Administration), 2016. National Water Model: Improving NOAA's Water Prediction Services.

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

Oktay, O., Schlemper, J., Le Folgoc, 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. In: Medical Imaging and Deep Learning Conference (MIDL 2018), July 4-6, 2018, Amsterdam, Netherlands. Last accessed April 17, 2020 at. https://arxiv.org/abs/1804.03999.

Passalacqua, P., Belmont, P., Foufoula-Georgiou, E., 2012. Automatic geomorphic feature extraction from lidar in flat and engineered landscapes. Water Resour. Res. 48, 1–18. https://doi.org/10.1029/2011WR010958.

Passalacqua, P., Trung, T.D., Foufoula-Georgiou, E., Sapiro, G., Dietrich, W.E., 2010. A geometric framework for channel network extraction from lidar: nonlinear diffusion and geodesic paths. J. Geophys. Res. 115, F)1002. https://doi.org/10.1029/2009JF001254.

Poppenga, S.K., Gesch, D.B., Worstell, B.B., 2013. Hydrography change detection: the usefulness of surface channels derived from LiDAR DEMS for updating mapped

hydrography. J. Am. Water Resour. Assoc. 49 (2), 371–389. https://doi.org/10.1111/jawr.12027.
Rahimi, A., Recht, B., 2008. Random features for large-scale kernel machines. In: Advances in Neural Information Processing Systems, pp. 1177–1184.
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566 (7743), 195–204.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, pp. 234–241.
Russell, P., 2008. Mapping headwater streams: intermittent and perennial headwater stream model development and spatial application. In: Final Report for Federal Highway Administration Contract Feasibility Study WBS: 36486.4.2. January 28. North Carolina Division of Water Quality, Raleigh.
Sangireddy, H., Stark, C.P., Kladzyk, A., Passalacqua, P., 2016. GeoNet: an open source software for the automatic and objective extraction of channel heads, channel network, and channel morphology from high resolution topography data. Environ. Model. Software 83, 58–73. ISSN 1364-8152.
Schmidhuber, J., 2015. Deep learning in neural networks: an overview. Neural Network. 61, 85–117.
Schultz, L.D., Heck, M.P., Hockman-Wert, D., Allai, T., Wenger, S., Cook, N.A., Dunham, J.B., 2017. Spatial and temporal variability in the effects of wildfire and drought on thermal habitat for a desert trout. J. Arid Environ. 145, 60–68.
Sharma, R., Xu, Z., Sugumaran, R., Oliveira, S., 2016. Parallel landscape driven data reduction & spatial interpolation algorithm for big LiDAR data. ISPRS Int. J. Geo-Inf. 5 (6), 97.
Shavers, E., Stanislawski, L., 2018. Streams do work: measuring the work of low-order streams on the landscape using point clouds. International Archives of the Photogrammetry. Remote Sensing & Spatial Information Sciences.
Sheng, J., Wilson, J.P., Chen, N., Devinny, J.S., Sayre, J.M., 2007. Evaluating the quality of the National Hydrography Dataset for watershed assessments in metropolitan regions. GIScience Remote Sens. 44 (3), 283–304.
Simley, J.D., Carswell Jr., W.J., 2009. The national map—Hydrography. US Geological Survey Fact Sheet 3054 (4).
Stanislawski, L.V., 2009. Feature pruning by upstream drainage area to support automated generalization of the United States National Hydrography Dataset. Comput. Environ. Urban Syst. 33 (5), 325–333.
Stanislawski, L.V., Survila, K., Wendel, J., Liu, L., Buttenfield, B.P., 2018. An open source high-performance solution to extract surface water drainage networks from diverse terrain conditions. Cartogr. Geogr. Inf. Sci. 45 (4), 319–328. https://doi.org/10.1080/15230406.2017.1337524.
Stoker, J.M., Abdullah, Q.A., Nayegandhi, A., Winehouse, J., 2016. Evaluation of single photon and Geiger mode lidar for the 3D Elevation Program. Rem. Sens. 8 (767) https://doi.org/10.3390/rs8090767.
Sun, Y., Zhang, X., Xin, Q., Huang, J., 2018. Developing a multi-filter convolutional neural network for semantic segmentation using high-resolution aerial imagery and LiDAR data. ISPRS J. Photogrammetry Remote Sens. 143, 3–14.
Tarboton, D.G., Bras, R.L., Rodriguez-Iturbe, I., 1991. On the extraction of channel networks from digital elevation data. Hydrol. Process. 5, 81–100. https://doi.org/10.1002/(ISSN)1099-1085.
Terziotti, S., Adkins, K., Aichele, S., Anderson, R., Archuleta, C., 2018. Testing the waters: integrating hydrography and elevation in national hydrography mapping. AWRA Water Resources IMPACT 20 (2), 28–29.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention is all you need. In: Advances in Neural Information Processing Systems, pp. 5998–6008.
Wang, S., 2010. A cyberGIS framework for the synthesis of cyberinfrastructure, GIS, and spatial analysis. Ann. Assoc. Am. Geogr. 100 (3), 535–557.
Wang, S., Goodchild, M.F., 2019. CyberGIS for Geospatial Innovation and Discovery. Springer, Dordrecht, Netherlands. https://doi.org/10.1007/978-94-024-1531-5.
Wang, S., Liu, Y., Padmanabhan, A., 2016. Open cyberGIS software for geospatial research and education in the big data era. SoftwareX 5, 1–5.
Wright, W., Nielsen, B., Mullen, J.D., Dowd, J., 2012. Agricultural Groundwater Policy during Drought: a Spatially Differentiated Approach for the Flint River Basin (No. 323-2016-11832.
Xu, Z., Guan, K., Casler, N., Peng, B., Wang, S., 2018. A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery. ISPRS J. Photogrammetry Remote Sens. 144, 423–434.
Xu, Z., Mountrakis, G., Quackenbush, L.J., 2017. Impervious surface extraction in imbalanced datasets: integrating partial results and multi-temporal information in an iterative one-class classifier. Int. J. Rem. Sens. 38 (1), 43–63.
Zhang, Z., 2018. Improved Adam optimizer for deep neural networks. In: 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE, pp. 1–2.
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 Geoscience and Remote Sensing Magazine 5 (4), 8–36.