



## Mapping stone walls in Northeastern USA using deep learning and LiDAR data

Ji Won Suh & William Ouimet

**To cite this article:** Ji Won Suh & William Ouimet (2023) Mapping stone walls in Northeastern USA using deep learning and LiDAR data, GIScience & Remote Sensing, 60:1, 2196117, DOI: [10.1080/15481603.2023.2196117](https://doi.org/10.1080/15481603.2023.2196117)

**To link to this article:** <https://doi.org/10.1080/15481603.2023.2196117>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 06 Apr 2023.



Submit your article to this journal [↗](#)



Article views: 428



View related articles [↗](#)



View Crossmark data [↗](#)

# Mapping stone walls in Northeastern USA using deep learning and LiDAR data

Ji Won Suh<sup>a,b</sup> and William Ouimet<sup>b,c</sup>

<sup>a</sup>Department of Natural Resources and the Environment, University of Connecticut, Storrs, CT, USA; <sup>b</sup>Department of Geography, University of Connecticut, Storrs, CT, USA; <sup>c</sup>Department of Earth Science, University of Connecticut, Storrs, CT, USA

## ABSTRACT

Stone walls are widespread and iconic landforms found throughout forested terrain in the Northeastern USA that were built during the 17th to early 20th centuries to delineate property boundaries and the edges of agricultural fields and pastures. As linear, or broadly curved, features that are typically >0.5 m high, 1–2 m wide, and >4–8 m long, stone walls are highly visible in LiDAR data, and mapping them is of broad interest to the cultural heritage sector as well as to researchers specifically focused on historic landscape reconstruction. However, existing mapping attempts have commonly relied on field surveys and manual digitization, which is time-consuming, especially when trying to complete mapping at broader scales. In response to this limitation, this study: (1) presents a novel framework to automate stone wall mapping using Deep Convolutional Neural Networks (DCNN) models (U-Net and ResUnet) and high-resolution airborne LiDAR, (2) evaluates model performance in two test sites against field verified stone walls, (3) investigates the factors that can influence model performance in terms of the quality of LiDAR data (e.g. ground point spacing), and (4) suggests post-processing for town-level mapping of stone walls (~120 km<sup>2</sup>). Both models performed well with respect to the Matthews Correlation Coefficient (MCC) score. U-Net scenario 3 achieved an MCC score of 0.87 at test site 1, while ResUnet scenario 3 (S3) had an MCC score of 0.80 at test site 2. In town-level test site 3, ResUnet S3 achieved the best F<sub>1</sub> score of 82% after post-processing. This study demonstrates the potential of automated mapping of anthropogenic features using our models.

## ARTICLE HISTORY

Received 19 December 2022  
Accepted 23 March 2023

## KEYWORDS

Automated anthropogenic feature mapping; stone walls; airborne LiDAR; deep convolutional neural networks; human impacts

## 1. Introduction

Stone walls are ubiquitous anthropogenic landforms found throughout forested terrain in the Northeastern USA that represent the legacy of historic, 17th- to early-20th-century land use practices in the region. The most common type of stone wall, fieldstone walls, are linear, or broadly curved, features that are composed of stones, typically >0.5 m high, 1–2 m wide and continuous for >4–8 m in length (Thorson 2023). The hundreds of thousands of fieldstone walls built across the Northeastern USA are a unique landscape feature that is the result of the region's glacial and land use history (Thorson 2002; Wessels 1997). Following European settlement, widespread forest clearing commonly exposed stones and sediment that had been deposited during deglaciation in the region between 17,000 and 21,000 years ago (Ridge 2004). Fieldstone walls were built through time as stones were moved to property boundaries and the edges of agricultural fields and pastures during the initial process of land clearing as well as the continued use of the land in the subsequent years, decade and even

centuries. The vast majority of these land use practices and walls were abandoned during mid-19th to early 20th centuries and reforestation followed (Foster et al. 2008; Johnson et al. 2021). Overall, stone walls are by far the most widespread and resilient anthropogenic landform representing previously deforested land throughout the previously glaciated Northeastern USA. Stone wall maps, as a proxy of historic agricultural activity, have been used to investigate historical land use dynamics and forest cover extent prior to mid-20th century (Johnson and Ouimet 2021, 2016, 2014).

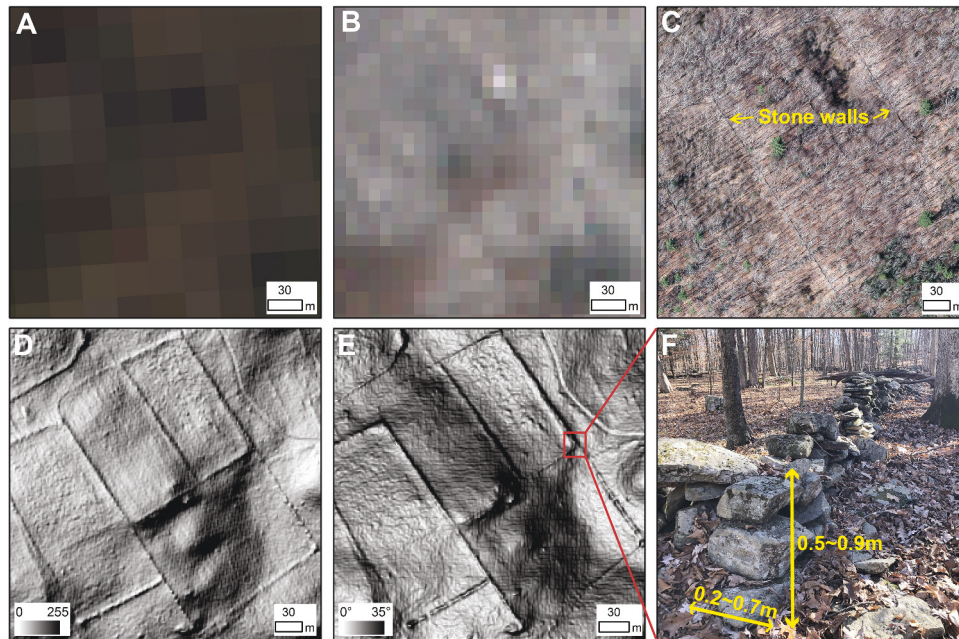
Mapping stone walls and quantifying the legacy of historical land use in the Northeastern USA is not possible through medium-resolution satellite imagery (e.g. Landsat 8 OLI and Sentinel 2) (Figure 1(a,b)) or high-resolution aerial imagery (Figure 1(c)) because stone walls are relatively small (i.e. less than 1–2 meters in width) and commonly rest under canopy cover. Airborne Light Detection and Ranging (LiDAR) solves these limitations (Figure 1(d,e)) (Risbøl and Gustavsen 2018; Doneus et al. 2008; Gallwey et al.

**CONTACT** Ji Won Suh  [ji.suh@uconn.edu](mailto:ji.suh@uconn.edu)

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.





**Figure 1.** Visibility of stone walls with different imagery sources. a: Landsat 8 OLI (30 m), 3 January 2016; b: Sentinel-2 (10 m), 18 February 2016; c: Aerial photography (0.1m), March 2016; d: LiDAR derived hillshade (1m, Azimuth 315°), e: LiDAR derived slope (1m); f: Stone walls in field (yellow refers to overall physical properties of stone walls (Johnson and Ouimet 2016)).

2019) because LiDAR can penetrate tree branches, shrubs, and underbrush and typically has 4–20 returns per  $\text{m}^2$ , allowing for the creation of Digital Elevation Models (DEMs) with a spatial resolution of  $\leq 1$  m. A diverse range of LiDAR DEM raster derivatives such as hillshade, slope, openness, Visualization for Archaeological Topography (VAT) and principal component analysis (PCA) can then be utilized for visualizing small anthropogenic features such as stone walls, charcoal hearths, grave mounds, and pits (Bennett et al. 2012; Hesse 2010; Doneus 2013; Evans et al. 2013; Štular et al. 2012; Howey et al. 2016; Kokalj, Zaksek, and Oštir 2013; Verbovšek, Popit, and Kokalj 2019; Suh et al. 2021; Johnson and Ouimet 2014). Utilizing these LiDAR derivatives and visualization techniques, manual digitization has been implemented as the most common approach to identify and map stone walls in the Northeastern USA (Johnson and Ouimet 2018, 2016, 2014). However, this approach is time-consuming, especially when trying to complete mapping over vast areas of land. Another approach is to use public crowd-sourced digitization via an ArcGIS web app covering the entire state (CT stone wall mapper<sup>1</sup> and NH stone wall mapper<sup>2</sup>). Even though the new mapper platform allows for crowd-sourced manual digitization over a broad area, manual digitization is still a time-inefficient approach and

an error can be introduced depending on the experience of the analysts (Leonard, Ouimet, and Dow 2021).

To overcome this inefficiency issue regarding mapping anthropogenic features and take advantage of the widespread availability of LiDAR datasets around the region (Johnson and Ouimet 2018, 2014), object-based image analysis (OBIA) has been applied for the detection and segmentation of anthropogenic features based on LiDAR derivatives (Niculiță 2020; Witharana, Ouimet, and Johnson 2018). OBIA is a semi-automated method including two steps: (1) extracting morphological characteristic (e.g. shape, size, and texture) of the target feature to differentiate it from the background or other objects, and (2) adjusting threshold to find the best parameters to detect target-like feature (Blaschke et al. 2000). A number of studies have utilized this approach using LiDAR products to extract, characterize, and classify various features such as relict charcoal hearths and kilns (Witharana, Ouimet, and Johnson 2018; Schneider et al. 2015; Suh et al. 2021), shell craters (Magnini, Bettineschi, and De Guio 2017), and mounds (Trier, Zortea, and Tonning 2015; Niculiță 2020; Davis, Lipo, and Sanger 2019; Freeland et al. 2016; Orenge et al. 2020). Previous studies have focused on features

that are isolated with circular shapes rather than those having continuous and linear shapes like a stone wall. While OBIA can be a useful and efficient mapping approach compared to manual digitization, it still requires significant time and effort to build a rule set based on morphological characteristics of the target feature and optimizing thresholding parameters.

As a fully automated approach, Deep Convolutional Neural Networks (DCNNs), also known as deep learning (DL), have been implemented in semantic segmentation and object detection using remote sensing imagery (Li et al. 2020; Zhong, Hu, and Zhou 2019; Waldner and Diakogiannis 2020). Semantic segmentation involves the classification of pixels according to semantic categories, which requires distinguishing boundaries between semantic categories, in addition to pixel-level classification. Different types of DCNN architectures have been developed (Ronneberger, Fischer, and Brox 2015; Zhou et al. 2018; Diakogiannis et al. 2020; Chen et al. 2018). Object detection, on the other hand, is the task of locating and identifying multiple objects in an image. The Region-based CNN (R-CNN) is one of DCNNs-based object detection models (Girshick et al. 2013). Unlike semantic segmentation model, the R-CNN-based model extracts a set of regions of interest (ROIs: region proposals) within an image to detect target objects, rather than processing the entire image. Then, the R-CNN model uses CNNs to extract features from the ROIs and classify the objects within the ROIs (He et al. 2020). These DCNN models are end-to-end process that automate the entire process of converting raw input data into the desired prediction without human intervention. With this advantage, its application is diverse in terms of image datasets and target mapping features. For example, multispectral satellite imagery (Zhong, Hu, and Zhou 2019; Li et al. 2020; Waldner and Diakogiannis 2020; Kussul et al. 2017) as well as aerial imagery (Sylvain, Drolet, and Brown 2019; Zhang et al. 2020) has been used for land cover (Mboga et al. 2020; Sylvain, Drolet, and Brown 2019; Srivastava, Vargas-Muñoz, and Tuia 2019; Li et al. 2020; Zhang et al. 2020; Rodríguez, Vitrià, and Mora 2020; Kim et al. 2018) and urban structures mapping (Zhang, Liu, and Wang 2018; Wang and Li 2019; Tan, Xiong, and Li 2018;

Abdollahi et al. 2020; Zhang et al. 2018). In geoscience field, imagery has been employed for landslide (Qi et al. 2020; Chen et al. 2020; Olteanu-Raimond et al. 2020) and fluvial or glacial lake mapping (Carbonneau et al. 2020; Wu et al. 2020).

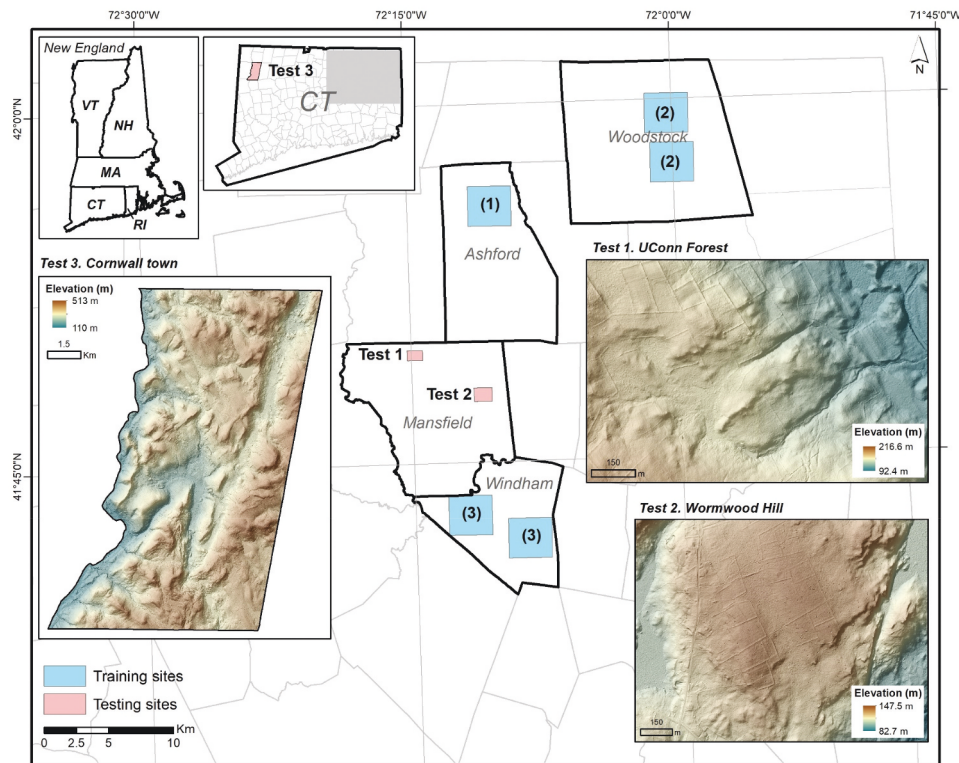
A few studies have been conducted to integrate LiDAR derivatives into DL models to identify archaeological objects (Trier, Reksten, and Løseth 2021; Verschoof-van der Vaart and Lambers 2019; Verschoof-Van Der Vaart et al. 2020). These studies utilized R-CNN-based models such as Faster R-CNN (Ren et al. 2017), to detect circular-shaped features such as charcoal kilns, pitfall traps, and grave mounds. However, there are several gaps in previous research: (1) there has been a lack of emphasis on the automated mapping of connected linear features, such as stone walls and historical roads (Verschoof-van der Vaart and Landauer 2021), (2) the application of semantic segmentation models such as U-Net has been limited in identifying small archaeological features from LiDAR derivatives (Suh et al. 2021), and (3) more research is needed to investigate the limitation of LiDAR-based mapping using DCNNs due to variations in LiDAR quality when applied to large sites (e.g.  $>100 \text{ km}^2$ ) and to mitigate these limitations by using additional reference data (e.g. land cover map) or GIS processing (Verschoof van der Vaart et al. 2022). Therefore, this study presents an innovative framework for the automated mapping of small-scale linear anthropogenic features, stone walls, in LiDAR derivatives up to town-level scale. In order to achieve this goal, two semantic segmentation DCNN models, U-Net and ResUnet, has been trained with different combinations of LiDAR derivatives from scratch. The trained models were first evaluated at two test sites (each with an area less than  $1.5 \text{ km}^2$ ) based on a field-verified stone wall map. Next, the potential factors that can affect model performance, such as the nature and the quality of LiDAR imagery, were investigated. Lastly, the trained models were applied to town-level stone wall mapping by combining GIS processing algorithms for post-processing to mitigate negative factors that may impact model performance. A further aim of this research is to use open-source statewide LiDAR data, making this study expandable to statewide stone wall mapping by other researchers, thereby enabling the reconstruction of historic deforestation and agricultural practices (Johnson and Ouimet 2021; Johnson et al. 2021).

## 2. Study sites and data

### 2.1. Study sites

Study sites are located in Connecticut, a state in the Northeastern USA (Figure 2), where stone walls are widely and densely distributed. Although historic stone walls are best preserved in the forested areas, they can be found all over the region in a variety of different land cover types, including deciduous/conifer forest, farmland, and developed/suburban region. Training sites were selected to cover a range of land cover types in order to train the model to deal with identifying stone walls in a variety of landscape types and conditions. Training sites specifically include dense deciduous/conifer forest areas, which account for over 80% (e.g. Ashford town) (Figure 2(1)), moderate to dense deciduous forest cover ranging from 62% to 70% (e.g. Woodstock town), farmland (e.g. portions of Woodstock town) (Figure 2(2)), and developed/suburban areas (e.g. portions of Windham town) (Figure 2(3)). Each training site covers an area of 13.5 km<sup>2</sup>.

To assess the performance of the model, three test sites were selected, two of which were small in extent and located in Mansfield town, CT (referred to as test site 1 and test site 2), and the third was town-level and located in Cornwall town, CT (referred to as test site 3) (Figure 2). Test site 1 is the northern portion of the Fenton Tract in UConn Forest (UF, 0.92 km<sup>2</sup>). The site primarily consists of dense deciduous forest cover, with over 90% coverage. Test site 2 is located along Wormwood Hill Road (WH, 1.46 km<sup>2</sup>). This site also consists mainly of dense deciduous forest cover, with over 78% coverage, and includes several structures such as road and buildings at the west part. Two of the test sites were chosen to evaluate the model's accuracy based on field-verified stone wall. Due to physical and personal property issue, we were able to conduct field-mapping in a limited area, with a focus on forested sites. Unlike two sites, test site 3, located in the Northeastern part of Connecticut (120 km<sup>2</sup>), was selected as the town-level evaluation site based on two factors: (1) a high density of stone walls and (2) a forest-dominant region, as the majority of



**Figure 2.** Study sites of this study with five training sites (including validation sites) and three test sites in Connecticut, USA. Training sites consist of (1) dense forested landscape in Ashford and (2) moderate to dense forested landscape and farmland in Woodstock, and (3) developed urban and suburban areas in Windham. Note that validation data is derived from training sites (See details about validation data in section 3.1.2). Test sites are (1) UConn forest (UF), (2) Wormwood Hill (WH) in Mansfield town, and (3) Cornwall town.



stone walls are typically found under forest cover. According to the 2015 land cover map (Connecticut Center for Land Use Education & Research (CT CLEAR) 2016), the land cover of Cornwall town is composed of deciduous forest (56.6%), conifer forest (23.2%), agricultural field (9.7%), developed area (1.5%), and other (9%). Cornwall was deemed an appropriate site to assess the performance of our models in mapping stone walls at a town level. The result of test site 1 and site 2 will be addressed in section 4.1 to 4.3 and that of test site 3 will be address in section 4.4.

## 2.2. Data

Data were collected for three main purposes: (1) to prepare training/validation datasets, (2) to build a reference dataset, and (3) to conduct post-processing (Table 1). For the training dataset, we used LiDAR-derived DEMs provided by CT ECO (2016). According to its metadata (Capitol Region Council of Governments (CRCoG) 2016b), the original point cloud had an average point spacing of 2 points per square meter, and the minimum spatial resolution of the LiDAR specification (Quality Level 2; QL2) was recommended to be 1 meter (Heidemann 2018). Based on this, CT ECO produced the 1 m resolution DEMs from the ground-classified points acquired in the spring of 2016 and we downloaded 1 m DEMs from CT ECO. Then, hillshade maps and slope maps were created for the training dataset. However, canopy cover such as conifers or underbrush can affect the quality of DEMs and derivatives locally because these areas may have lower point densities of ground-classified points, or, in the case of dense underbrush, points classified as ground may not actually represent the ground surface, leading to greater

elevation error in the resultant DEM. We will discuss how these LiDAR quality issues affect DCNN model performance when mapping anthropogenic landform in section 4.3.4 in detail. For the reference stone wall dataset, we edited a previously digitized stone walls shapefile by Johnson and Ouimet (2014) which was based on 2011 LiDAR dataset rather than 2016 LiDAR dataset used in this study. Our update was performed using hillshade maps, high-resolution leaf-off aerial orthophoto, and Google Street View images to confirm the existence of stone walls. Stone wall digitization was primarily undertaken using a LiDAR-derived hillshade map since the feature is more distinguishable than in a slope map and leaf-off orthophotography. Another stone wall polyline dataset was built by field mapping (referred to as field verified stone walls) particularly in test site 1 and 2. For test site 3, 2015 land cover map (30 m) and impervious shapefile was used to conduct post-processing of stone wall map predicted by model. Detailed information about post-processing will be addressed in section 3.4.

## 3. Methodology

We proposed a workflow for a small-scale anthropogenic feature mapping using LiDAR derivatives and DCNN models (i.e. U-Net and ResUnet) (Figure 3). Details for each step will be addressed in the following subsections.

### 3.1. Preparing the image and reference datasets

#### 3.1.1. LiDAR derivatives preparation

Given that DCNN is a supervised model, the first step is to prepare a training/validation dataset for the model training by combining the LiDAR derivatives with the rasterized reference data of the stone wall image. For the

**Table 1.** List of collected data.

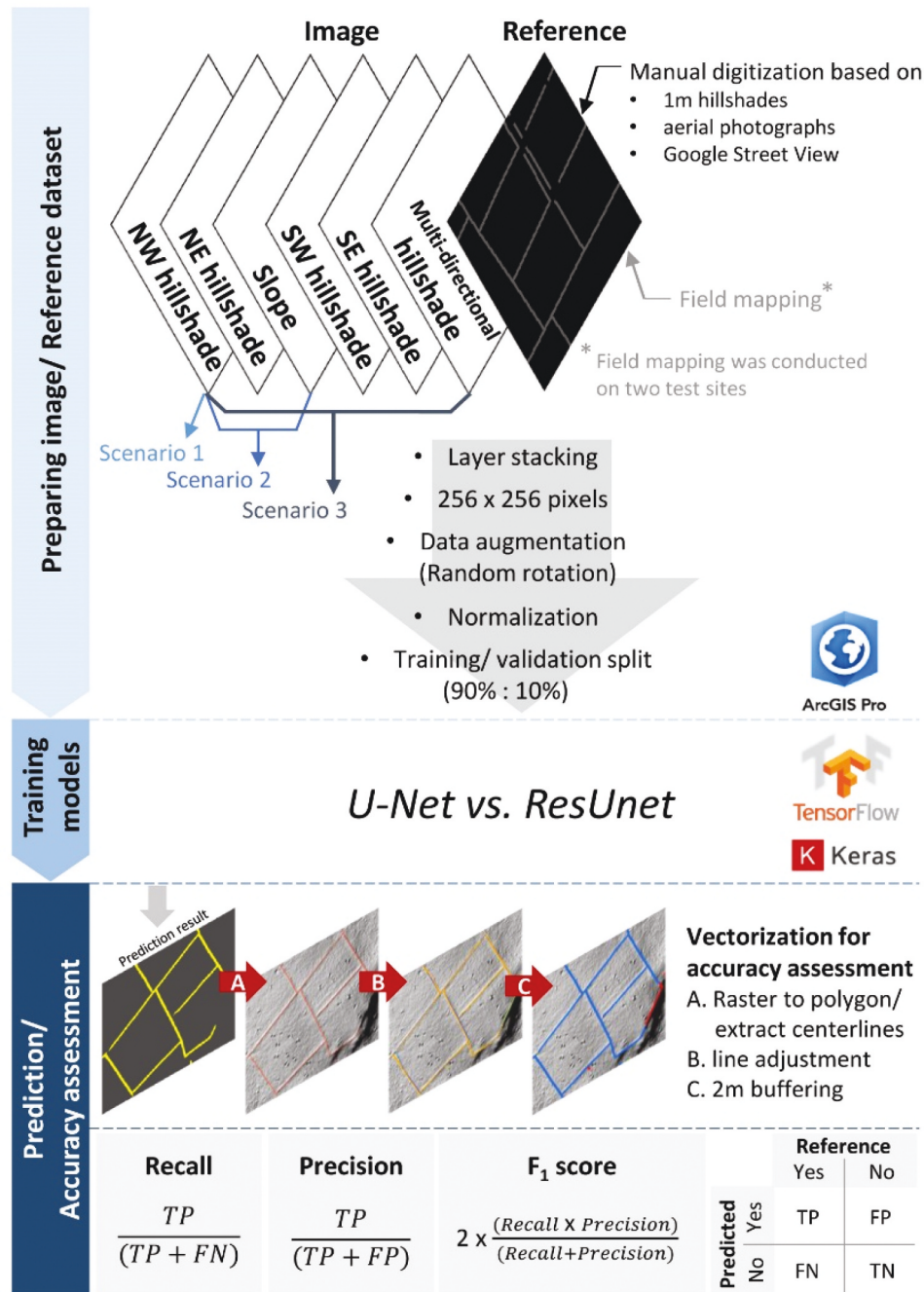
| Data                              | Resolution    | Source                    | Purpose                                                                 |
|-----------------------------------|---------------|---------------------------|-------------------------------------------------------------------------|
| 2016 LiDAR-derived DEMs           | 1m            | CT ECO*                   | Training/validation dataset                                             |
| Hillshade map                     | 1m            | LiDAR derivatives         | Training/validation dataset and reference dataset (manual digitization) |
| Slope map                         | 1m            | LiDAR derivatives         | Training/validation dataset                                             |
| 2016 Orthophoto (leaf-off)        | 3 bands, 0.1m | CRCoG (2016a)**           | Reference dataset (Supplement for manual digitization)                  |
| Google Street View                | -             | Google Map                | Reference dataset (Supplement for manual digitization)                  |
| Digitized stone walls***          | -             | Johnson and Ouimet (2014) | Training/validation dataset and reference dataset                       |
| Field verified stone walls        | -             | Field work                | Reference dataset                                                       |
| 2015 Land cover map               | 30m           | CT CLEAR (2016)****       | Post-processing                                                         |
| 2012 Impervious cover (shapefile) |               | CT CLEAR (2017)           | Post-processing                                                         |

\*Connecticut Environmental Conditions Online (CT ECO).

\*\*Capitol Region Council of Governments (CRCoG).

\*\*\*Digitized shapefiles are available via <https://connecticut.maps.arcgis.com/home/item.html?id=0d956858d17a4366a630eded3c2ff703> (accessed on 2 November 2022).

\*\*\*\*Center for Land Use Education & Research (CT CLEAR).



**Figure 3.** Workflow for the stone wall detection. Abbreviation: TP (true positives), FP (false positives), FN (false negatives), TN (true negatives).

LiDAR derivatives, we used three different scenarios (S1, S2, and S3) depending on the types of input rasters (i.e. LiDAR-derived hillshades or slope map), different sunlight angle of hillshade images (i.e. NW, NE, SW, SE, and multi-direction), and the number of input rasters (i.e. 1, 3, and 6) (Table 2). All hillshade images were created using the Hillshade tool in ArcGIS Pro with a z-factor of 3.5, and slope images (degrees) were created using Slope tool with a z-factor of 3, then converted to an 8-bit image

with values ranging from 0 to 255. We assumed different sunlight azimuth angle plays an important role in improving model performance since it can visualize and highlight morphological property of stone wall feature from a different perspective.

### 3.1.2. Reference preparation

As briefly stated in section 2.2, reference datasets were constructed by manual digitization and field

**Table 2.** Specific information of three input scenarios used in this study.

| Scenarios       | Types of input Rasters                | # of Rasters |
|-----------------|---------------------------------------|--------------|
| Scenario 1 (S1) | NE hillshade (sunlight azimuth: 315°) | 1            |
| Scenario 2 (S2) | NE hillshade (315°)                   | 3            |
|                 | NW hillshade (45°)                    |              |
| Scenario 3 (S3) | Slope                                 | 6            |
|                 | NE hillshade (315°)                   |              |
|                 | NW hillshade (45°)                    |              |
|                 | SE hillshade (135°)                   |              |
|                 | SW hillshade (225°)                   |              |
|                 | Multi-directional hillshade           |              |
|                 | Slope                                 |              |

mapping. Even though we used the digitized stone walls by Johnson and Ouimet (2014), extra editing was required to examine topographic features that were newly identifiable in the updated 2016 LiDAR hillshade map. Primarily, we manually digitized stone wall features over 1 m NE/NW hillshade maps using ArcGIS Pro software. Leaf-off aerial photographs and Google Street View were used as supplementary imagery in cases where stone wall features were questionable due to areas of light or dark in the hillshade maps. In some cases, the street view map allowed us to identify stone walls clearly along roads that are not visible in hillshade maps. Next, field mapping was conducted to build an accurate reference dataset of stone walls within study sites to compare LiDAR-based model results against. Particularly, our field mapping was focused on two test sites (UF and WH) considering accessibility issue such as private property. Once the reference stone wall dataset was created as polyline, then 2 m buffering was conducted to capture the entire width of the stone wall feature topography and lastly, buffered polygons were rasterized.

### 3.1.3. Training/validation patches

With LiDAR derivatives and rasterized reference data, training/validation patches for two DCNN models were prepared by the following steps. (Figure 3). Firstly, training/validation data were created by stacking each scenario of LiDAR derivatives and reference raster, respectively. Secondly, to reduce computational burden, we clipped this composited training/validation data into the given size of patches (i.e. 256 by 256 pixels), using a stride of 64 pixels in x and y direction to obtain a large number of training patches. Thirdly, we applied data augmentation to the clipped patches using a random rotation method including  $-90^\circ$ ,  $90^\circ$ , and  $180^\circ$  rotation to increase the total number of input patches. With randomly

rotating input patches, it also allows for training model with various directional stone walls. As a result, the total number of input patches was 12,774. Next, we normalized the pixel value of the input patches to 0–1 to increase training speed and improve feature learning during the model training. Lastly, the total number of input patches were split into training dataset and validation dataset (90% vs. 10%); 11,496 and 1,278, respectively.

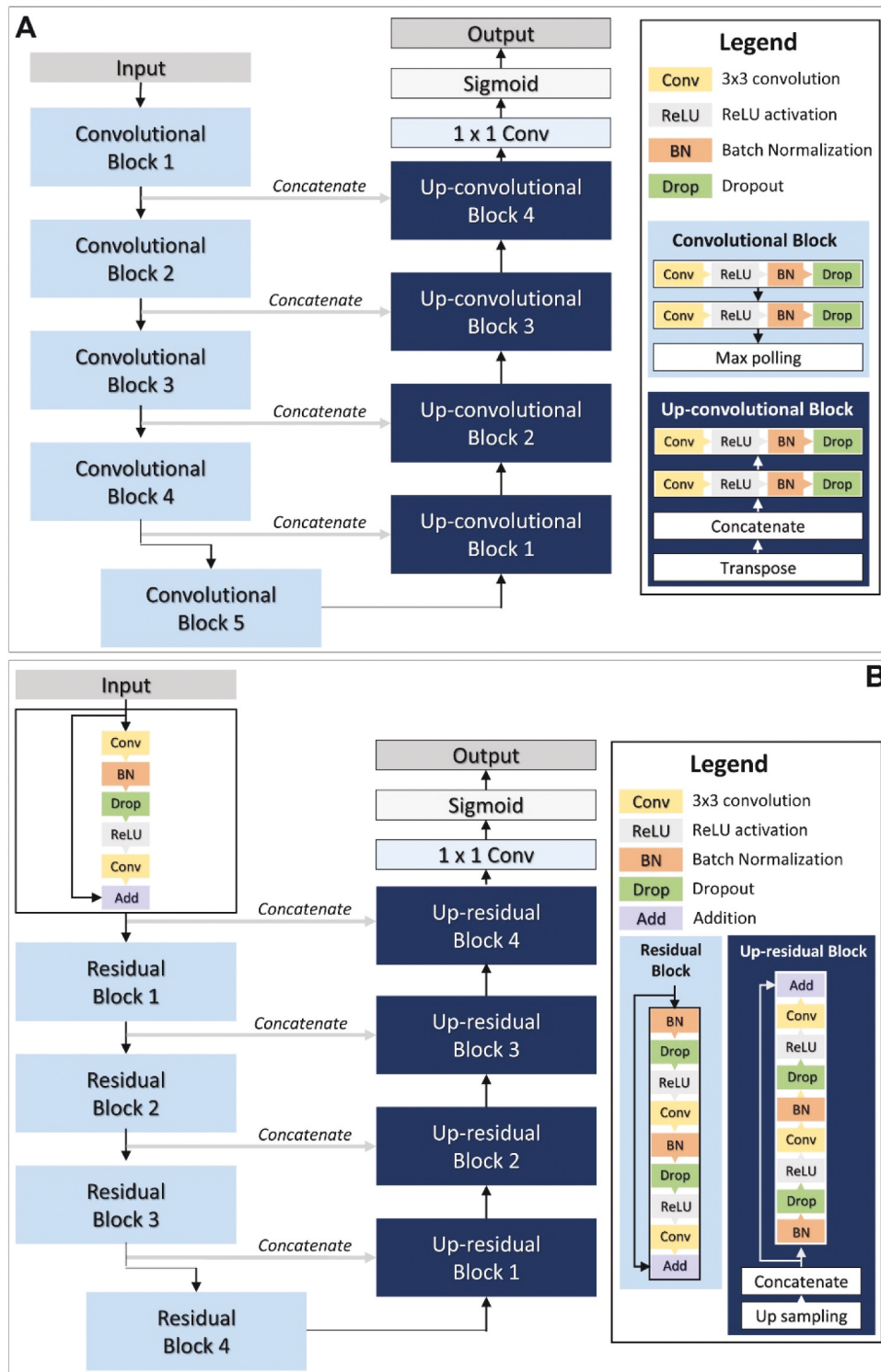
## 3.2. Training models

### 3.2.1. Model architectures

In previous research, the U-Net model has been shown to achieve highly accurate segmentation results even with small input datasets (Ronneberger, Fischer, and Brox 2015; Chen et al. 2022; Kugelman et al. 2022). Additionally, as we trained our model from scratch, we prepared our training dataset by manually digitizing stone walls. Manual digitization is a labor-intensive process, so we were only able to prepare a limited-size of training dataset. However, U-Net tends to show high performance with limited data. Therefore, we used two U-Net-based models, one is a modified U-Net from original U-Net architecture developed by Ronneberger, Fischer, and Brox (2015) and the other is a ResUnet, incorporating Residual Network (ResNet) (He et al. 2015) to U-Net architecture. Although U-Net was originally proposed for the semantic segmentation of panchromatic biomedical images, it has been applied to satellite imagery as well as aerial imagery in recent years (Mboga et al. 2020; Stoian et al. 2019; Yan et al. 2021; Waldner and Diakogiannis 2020).

The structure of the U-Net model used in this study consists of encoding and decoding branches, like U-shape (Figure 4(a)). In the encoding branch, the model extracts a given number of feature maps (here we used 32, 64, 128, 256, and 512 in





**Figure 4.** Model architectures used in this study (A: U-Net, B: ResUNet).

convolutional block 1, 2, 3, 4, and 5, respectively) from input patches by passing through convolutional blocks. Each convolutional block includes two sets of  $3 \times 3$  filter convolution layers, a Rectified Linear Unit (ReLU) activation layer, a batch normalization layer

and dropout layer, then it is connected to  $2 \times 2$  size of max pooling layer. In particular, we used batch normalization (Ioffe and Szegedy 2015) and dropout strategies to avoid an overfitting issue. While passing through convolutional blocks, the size of the input

feature map (i.e. length and width) is halved but the depth (i.e. the number of feature maps) is doubled. In the U-Net model, the lowest resolution of feature map is  $16 \times 16 \times 512$  (width x height x depth) in convolutional block 5. In the decoding branch, a series of up-convolutional blocks helps to restore spatial information that was lost during max pooling layer in the convolutional block. Particularly, it includes a transposed convolutional layer, concatenated layer, two sets of convolutional layers with ReLU activation layer and dropout layer. After going through four up-convolutional blocks, it passed  $1 \times 1$  convolution layer to reduce the number of channels or the depth of the feature map produced by the previous layer. In the last step, semantic segmentation is conducted by a sigmoid function to create a binary map (i.e. “stone wall” or “non stone wall”).

The second model, ResUnet, has similar architecture to the U-Net model except for adding a residual unit in each block (Figure 4(b)). ResUnet, proposed by (Zhang, Liu, and Wang 2018), was used as a baseline model. We modified the number of feature maps at each residual block, added one more residual block, as well as an up-residual block to be comparable to our U-Net model architecture. The concept of residual unit was first proposed by He et al. (2015) to address a vanishing gradient issue that occurs when the neural network is going deeper. The application of residual unit to the U-Net model shows improvement with respect to model performance in previous studies (Zhang, Liu, and Wang 2018; Liu et al. 2020).

Table 3 shows a summary of specific model parameters used in both U-Net and ResUNet models. We used input size (width and height) of 256 by 256 pixels and the number of input channels varied depending on the composition of bands scenarios. The number of feature maps for convolutional blocks

**Table 3.** Summary of model architecture for both U-Net and ResUNet.

| Parameter                              | Value                                              |
|----------------------------------------|----------------------------------------------------|
| Input size (width and height) (pixels) | 256 by 256                                         |
| Number of input channels               | S1: 1/S2: 3/S3: 6                                  |
| Number of training/validation patches  | Training: 11496 (90%)<br>Validation: 1,278 (10%)   |
| Number of feature maps                 | [32, 64, 128, 256, 512]                            |
| Convolution filter size (pixels)       | 3 by 3                                             |
| Activation function                    | ReLU/Sigmoid (only used right before output layer) |
| Total number of trainable parameters   | U-Net: 7,764,962<br>ResUNet: 18840,610             |

**Table 4.** Hyperparameters for model training.

| Hyperparameter     | Value                                   |
|--------------------|-----------------------------------------|
| Batch size (count) | 8                                       |
| Optimizer          | Adam                                    |
| Learning rate      | Initially starting from 0.001           |
| Loss function      | Binary Cross Entropy                    |
| Epochs             | Up to 30 (used early stopping callback) |

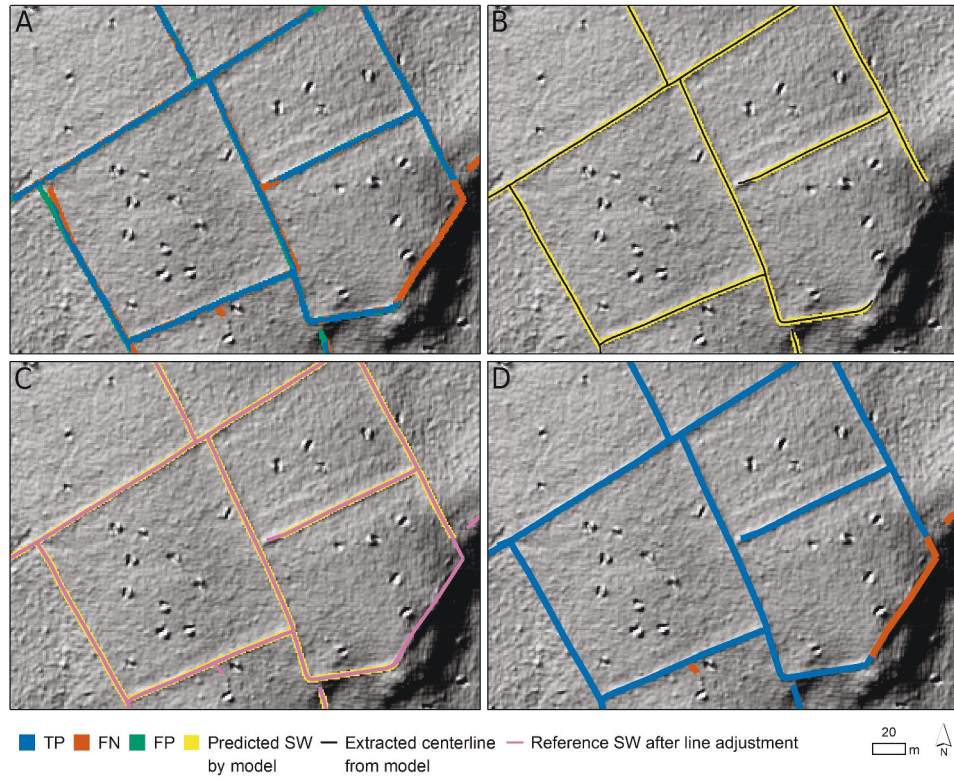
or residual blocks was set as 32, 64, 128, 256, and 512 as networks go deeper. For the activation function, ReLU (which has been widely used) was adopted, and the Sigmoid function was used before the output layer since target classes are binary. As a result, the total number of trainable parameters for U-Net and ResUNet was about 7 million and 18 million, respectively.

### 3.2.2. Model parameters/hyperparameters

Our model was trained on the single GPU, 8 G RTX 2070 (compute capability = 7.5) with Keras API of TensorFlow 2.1 version. During the training process, models were tuned by adjusting hyperparameters to avoid overfitting and to improve model performance. Table 4 shows the final hyperparameters. Due to an out of memory issue (OOM) in training the ResUNet model, the batch size for two models was set to 8, which takes more time on training compared to large batch size. Despite this fact, the large batch size does not always have a positive influence on model performance. Another important hyperparameter is a learning rate that can influence training efficiency. We used an Adam optimizer (Kingma and Ba 2015) (initial learning rate = 0.001) and it was multiplied by 0.1 whenever the validation loss value stopped diminishing during the previous three epochs. In addition, the binary cross-entropy loss function was used to monitor model performance which is suitable for binary classification with less calculation burden. Lastly, models were trained for 30 epochs. However, it recalled early stopping callback when validation loss would not improve during the previous four epochs.

### 3.3. Model prediction and vectorization of raster results

Through the training process of both U-Net and ResUnet models for three scenarios, we had six trained models and applied them to two test sites. Before evaluating the models' accuracy from the model prediction result, vectorization of the



**Figure 5.** Vectorization process of model prediction result to minimize a potential error in accuracy assessment. A: example of errors occurred in raster-based accuracy assessment result. B: predicted stone walls (SW) by model (yellow pixel) and its centerline (black line); C: reference stone wall after line adjustment (pink line) based on extracted centerline from model (black line); D: Accuracy assessment result after vectorization process. Note that TN is true negative, TP is true positive, FN is false negative, and FP is false positive.

prediction result was required to minimize error that regards true positive as false positive or false negative (Figure 5(A)). This error can be caused by the fact that two rasters (rasterized stone wall reference map and model prediction result) are not perfectly aligned even through the model identifies stone wall correctly. Given that this error can influence on accuracy result, we performed three-step processing to conduct a vector-based calculation.

- (1) Extracting a centerline (black line) from the vectorized model prediction result (yellow pixel) (Figure 5(B)).
- (2) Comparing this centerline to the reference stone wall polyline and removing the reference stone wall line if the centerline and the reference line are overlapped each other within 2 m buffering threshold (Figure 5(C)).
- (3) 2 m buffering of updated reference line and conducting accuracy assessment (Figure 5(D)).

### 3.4. Accuracy assessment

#### 3.4.1. Test site 1 and 2

Accuracy assessment was conducted to evaluate model performance based on the areas (in  $m^2$ ) of TP, FN, FP, and TN using the following evaluation metrics; recall, precision,  $F_1$ , and Matthews Correlation Coefficient (MCC) score (Equation (1)–(4)). For the test site 1 and 2, field-verified stone wall was used as a reference map when calculating accuracy assessment.

$$Recall = \frac{TP}{TP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$F_1score = 2 * \frac{(Recall * Precision)}{(Recall + Precision)} \quad (3)$$



$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4)$$

where TP means the feature regarded as a stone wall by the model is also a true stone wall in the reference data, FN is the case that model missed a stone wall that is actually a true stone wall in the reference data, and lastly, FP is the case that model detects feature as a stone wall that is not a stone wall in the reference data.

Recall measures how good the model finds all positives, all existing stone walls in this case (i.e. the sum of TP and FN) and it is calculated by Equation (1). Particularly, a FN is related to omission error so that recall can be defined as 1-omission error. Precision evaluates how accurate the model's prediction result (i.e. the sum of TP and FP) is. It is calculated by Equation (2), and it can be described as 1-commission error since FP is related to commission error. Based on the recall and precision, the  $F_1$  score is widely used to assess overall model performance. MCC is a binary classification accuracy metric that considers both the true positive rate (TPR) and the false positive rate (FPR). Unlike  $F_1$  score, MCC provides a balanced measure of the trade-off between TPR and FPR, which makes it a suitable for imbalanced datasets. MCC ranges from  $-1$  to  $1$ , where a value of  $1$  represents a perfect prediction,  $0$  represents a random prediction, and  $-1$  represents a completely incorrect prediction (Chicco and Jurman 2020). For test site 1 and 2, these four metrics were calculated by the area of buffered polygons from model prediction lines and updated reference lines after line adjustment explained in Figure 5.

### 3.4.2. Test site 3

For town-level mapping (test site 3), a larger test area includes a diverse array of terrain and land cover conditions, which could potentially weaken model performance as reported in previous studies (Trier, Reksten, and Løseth 2021; Suh et al. 2021; Verschoof van der Vaart et al. 2022). To address this issue, post-processing was conducted for test site 3 before accuracy assessment.

The workflow of post-processing and accuracy assessment of town-level model results is demonstrated in Figure 6. During the post-processing step, an impervious cover shapefile from 2012, provided by

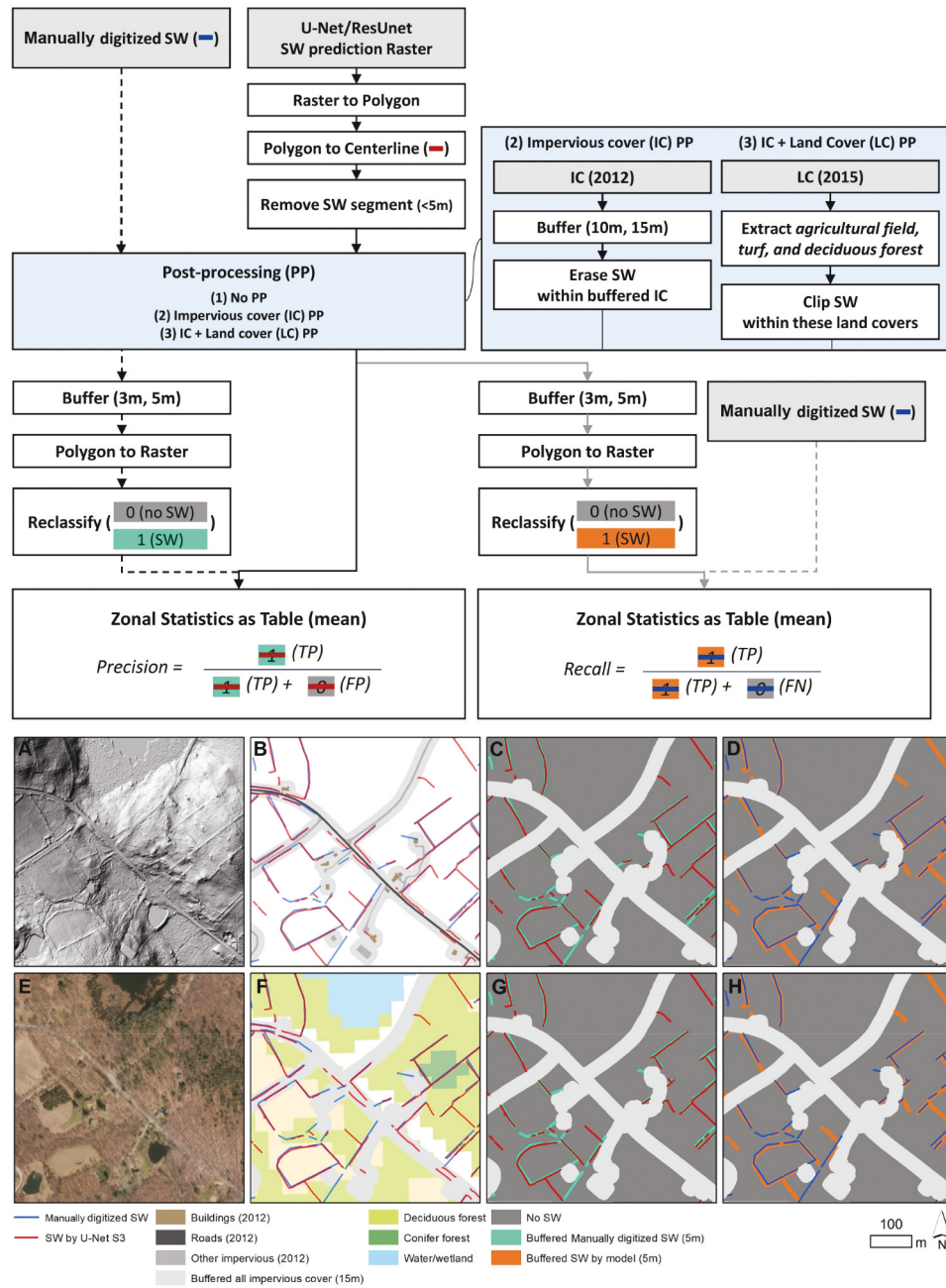
CT CLEAR (2017), was used to mask out stone walls near roads and building that can be confused with stone walls LiDAR derivatives due to similar morphology (this process is referred to as IC PP: impervious cover post-processing). The impervious cover shapefile includes three types of impervious covers: road, building, and other. All three types were buffered based on thresholds of 10 m and 15 m (Figure 6(b)) as most stone wall-like features (but not actual stone walls) near impervious cover were found within these buffer ranges. Next, the manually digitized stone wall and model-predicted stone wall shapefiles were removed from the buffered impervious cover using the Erase tool in ArcGIS Pro to exclude stone walls or stone wall-like road curbs or road edge segments (Figure 6(c,d)). As a reference map for accuracy assessment, the manually digitized stone wall based on hillshades was used.

In addition to IC PP, land cover was considered to exclude errors related to the low quality of LiDAR. The issue with LiDAR quality was also discussed in previous studies (Crow et al. 2007; Doneus et al. 2008; Suh et al. 2021; Verschoof van der Vaart et al. 2022) which found that a lack of ground-classified points due to conifer trees can lead to irregular interpolation of terrain surfaces, resulting in commission or omission errors. A recent land cover map from 2015 (spatial resolution: 30 m) provided by CT CLEAR (2016) was used to extract agricultural fields, turf, and deciduous forests as polygons where the quality of LiDAR is high (this process is referred to as IC LC PP: impervious cover and land cover post-processing) (Figure 6(f)). Then stone walls in these polygons were then extracted using the Clip tool in ArcGIS (Figure 6(g,h)).

These two post-processing methods, IC PP and IC LC PP, also contribute to controlling errors derived from digitizing users because reference data for accuracy assessment is based on a manually digitized shapefile rather than field mapping. As discussed in Leonard, Ouimet, and Dow (2021), users have different interpretations and errors associated with digitizing stone walls using LiDAR derivatives. Disagreement in digitizing stone walls between users is more likely to occur near stone-like features along roads and in areas of irregular morphology derived from the low point density of the LiDAR data.

As described in Figure 6, buffer thresholds of 3 m and 5 m were used to calculate precision and recall





**Figure 6.** Workflow of post-processing (PP) for accuracy assessment in town-level model prediction result. (a: NW hillshade map; b: Manually digitized stone wall (SW) and SW predicted by U-Net S3 model overlaying three types of impervious cover (IC) in 2012 and buffered all impervious cover (Data source: CT CLEAR (2017)); c: Precision example after impervious cover post-processing (IC PP); d: Recall example after impervious cover post-processing (IC PP); e: leaf-off aerial photography in 2016 (CRCoG 2016a); f: Manually digitized stone wall (SW) and SW predicted by U-Net S3 model overlaying land cover (LC) of 2015; g: Precision example after impervious and land cover post-processing (IC LC PP); h: Recall example after impervious and land cover post-processing (IC LC PP)).

scores, taking into account that the width of stone walls are described on hillshades in Cornwall can reach up to 10 m. The reason for using buffer thresholds is that manually digitized stone wall (vector line) may slightly differ from model-predicted stone wall (vector line). While the model tends to detect the centerline of stone walls in the hillshades, manual

digitizing results do not always focus on the center of stone walls in the hillshades. Cases where the stone wall polyline predicted by the model is aligned within a 3 m or 5 m buffered digitized stone wall polyline were considered TP; otherwise, they were considered FP in terms of precision (Figure 6(c,g)). In case of recall, cases where the digitized stone wall polyline

is aligned within 3 m or 5 m buffered stone wall poly-line predicted by model were considered TP; otherwise, they were considered FN (Figure 6(d,h)). As TN is not extracted in this method, it has not been possible to calculate the MCC score for test site 3.

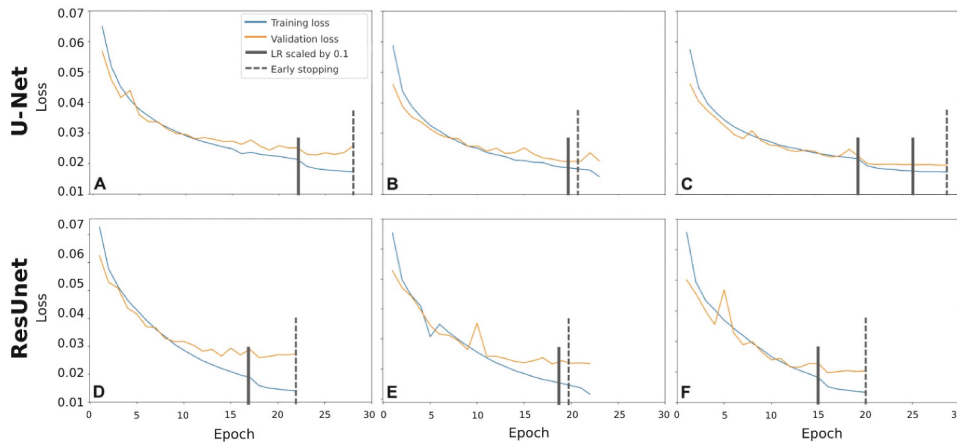
#### 4. Results and discussion

We trained both U-Net and ResUnet models for each scenario (S1, S2, and S3) and it took 200 seconds per epoch to train U-Net, and 620 seconds per epoch for ResUnet. As we can see through the difference between the number of trainable model parameters for each model, training speed of U-Net is three times faster than that of ResUnet. Figure 7 illustrates binary cross entropy losses (training/validation) for the U-Net/ResUnet models by each scenario. Minor difference between training and validation loss demonstrates that our models trained well without overfitting or underfitting issues. In addition, the training process was stable based on a smooth decreasing trend. As stated in Section 3.2.2, the learning rate was reduced when the model was not improved during the previous three epochs, which

are marked as black solid line in Figure 7. Except for S2, the decay of learning rate helped models to reduce loss value, although the gap between training loss and validation loss increases. In S2, the model reached minimum loss value before the learning rate was adjusted. For model prediction, we used the best model with the lowest validation loss value, and its accuracy assessment result will be addressed in the following sections.

##### 4.1. Accuracy assessment results for two test sites

We calculated the area of TP, FN (omission error), FP (commission error), and TN in two test sites. Table 5 shows the model performance results based on recall, precision,  $F_1$ , and MCC scores at different models and scenarios. Taking into consideration the balance between TP, FN, FP, and TN, our model achieved the best MCC score of 0.87 and 0.8 in two test sites, respectively. The MCC and  $F_1$  score indicated that S3 performed the best in both models at two test sites, which means there is a positive relationship between model performance and the number of input rasters (i.e. different directional hillshade maps and slope



**Figure 7.** Binary cross entropy loss curve per epoch for U-Net/ResUnet models and three scenarios (S1, S2, and S3). (A: U-Net S1, B: U-Net S2, C: U-Net S3, D: ResUnet S1, E: ResUnet S2, F ResUnet S3).

**Table 5.** Result of model performance with different models and scenarios.

| Site #         | Scenario | U-Net  |           |              |             | ResUnet |           |              |             |
|----------------|----------|--------|-----------|--------------|-------------|---------|-----------|--------------|-------------|
|                |          | Recall | Precision | $F_1$        | MCC         | Recall  | Precision | $F_1$        | MCC         |
| Test 1<br>(UF) | S1       | 70.5%  | 94.1%     | 80.6%        | 0.80        | 67.0%   | 97.3%     | 79.4%        | 0.80        |
|                | S2       | 84.0%  | 92.3%     | 88.0%        | 0.87        | 78.3%   | 92.0%     | 84.6%        | 0.83        |
|                | S3       | 83.5%  | 93.7%     | <b>88.3%</b> | <b>0.87</b> | 82.4%   | 94.0%     | 87.8%        | 0.87        |
| Test 2<br>(WH) | S1       | 60.0%  | 93.8%     | 73.2%        | 0.74        | 60.0%   | 94.2%     | 73.3%        | 0.74        |
|                | S2       | 64.4%  | 95.3%     | 76.9%        | 0.78        | 61.1%   | 96.7%     | 74.9%        | 0.76        |
|                | S3       | 64.7%  | 95.0%     | 76.9%        | 0.77        | 69.6%   | 94.9%     | <b>80.3%</b> | <b>0.80</b> |

map). By comparing the MCC and  $F_1$  score between U-Net and ResUnet, we found that the U-Net model outperformed ResUnet at test site 1, while ResUnet outperformed U-Net at test site 2.

In terms of recall and precision scores, both models showed high precision scores but relatively low recall scores. The difference between recall and precision scores reduced when more LiDAR-derived rasters were used. In particular, multi-layered hillshade and slope map (i.e. S2 and S3) tended to play a critical role in reducing FN, while they do not contribute to reducing FP. This result suggests that a diverse array of LiDAR derivatives allows for identifying stone walls with different physical properties (size, height, number of stacked stones, etc.) from normal stone walls.

## 4.2. Visualization of model prediction results for two test sites

### 4.2.1. Test site 1 (UF)

Figure 8 shows the spatial distribution of TP, FN, and FP cases in test site 1 by models and scenarios. The TP cases in both model scenarios are mainly distributed in areas with moderate to dense forest cover. The FN cases are mainly distributed along brooks or shaded surfaces in the hillshade maps. The FP cases are located on trails, paths, or wall-like features (such as a field edge with a pile of leaves) in the hillshade image due to morphological similarities. Figure 8 also visually demonstrates the improvement in model performance for each scenario. Many of the FN cases in S1 are distributed near NW-SE directional walls, but they tend to be classified as TP cases in S2 and S3. This is partly due to the fact that sunlight from the NW direction does not create shade or light for NW-SE directional stone walls in the NW hillshade image. This will be discussed in detail in Section 4.3.

### 4.2.2. Test site 2 (WH)

Figure 9 demonstrates the spatial distribution of TP, FN, and FP by models and scenarios at test site 2. Similar to test site 1, the TP cases are primarily observed in moderate to dense deciduous vegetation. The FN cases are found in forested areas for the similar reasons as test case 1, but they are more prevalent in the western part of the test site near roads and building foundations. This is likely due to the difficulties in distinguishing the morphological characteristics of stone walls, such as height, width, and length, from

other man-made structures in the hillshade or slope image. As our models were conservative in mapping stone walls (low recall), there are only a few instances of FP in test site 2, and these are distributed along roads or trail paths.

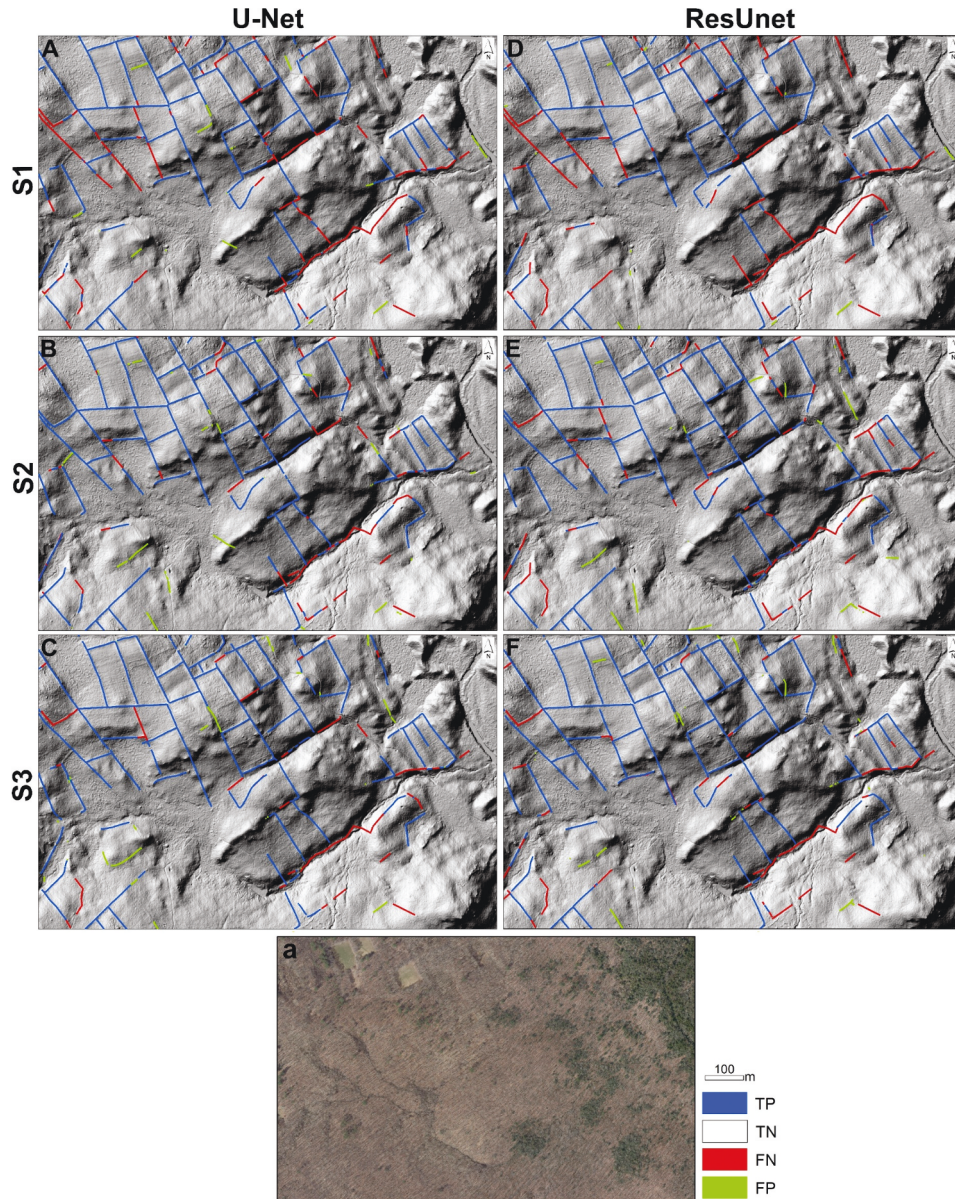
Based on the results of both test sites, we conclude that the models perform best at identifying stone walls in areas dominated by deciduous forest that are away from roads and buildings. This is likely because the morphology of stone walls (>0.5 m high, 1–2 m wide and continuous for >4–8 m in length (Thorson 2023) make them visible in hillshades or slope maps within deciduous forests or in open areas where LiDAR quality is high due to many points representing the ground surface being included in the point cloud (Doneus et al. 2008).

Conifer forests and areas with a low density of ground points in general, meanwhile, would likely lead to FN (omission error) because the derivative the LiDAR derived DEMs would be impacted by having too few points to interpolate between, making it difficult to identify small-scale anthropogenic features (Verschoof van der Vaart et al. 2022; Suh et al. 2021). In addition, we also found that the multi-layered input image (S3) contributes to reducing FN since the model learns various examples of stone wall features from visualizations of different directional hillshade rasters. However, having a large number of input channels does not necessarily enhance the model performance since each model (e.g. U-Net S1, U-Net S2, and U-Net S3) extracts distinct feature maps during the training process and predicts stone walls in the test sites based on these feature maps. Although input rasters of S3 include all rasters of S2, there were instances where stone wall segments were correctly identified by U-Net S2, while these segments were missed (omission error) by U-Net S3 (Figure 9). This highlights the drawbacks of DCNNs as black-box models, where the complexities involved in training the models are challenging to interpret.

## 4.3. Factors that can cause FP (commission error) and FN (omission error)

As shown in the results, multi-stacked LiDAR products contribute to identifying a small-scale morphological feature, such as a stone wall. However, the quality and nature of LiDAR products are essential factors that influence the model prediction results, given that LiDAR hillshades and slope





**Figure 8.** The distribution of TP, FN, FP, and TN results from different models and scenarios at test site 1 (UF). a: high-resolution orthophotography (CT ECO 2019); A: U-Net S1; B: U-Net S2; C: U-Net S3; D: ResUnet S1; E: ResUnet S2; F: ResUnet S3. NW hillshade map is a background. Note that TN symbol is transparent to visualize background hillshades.

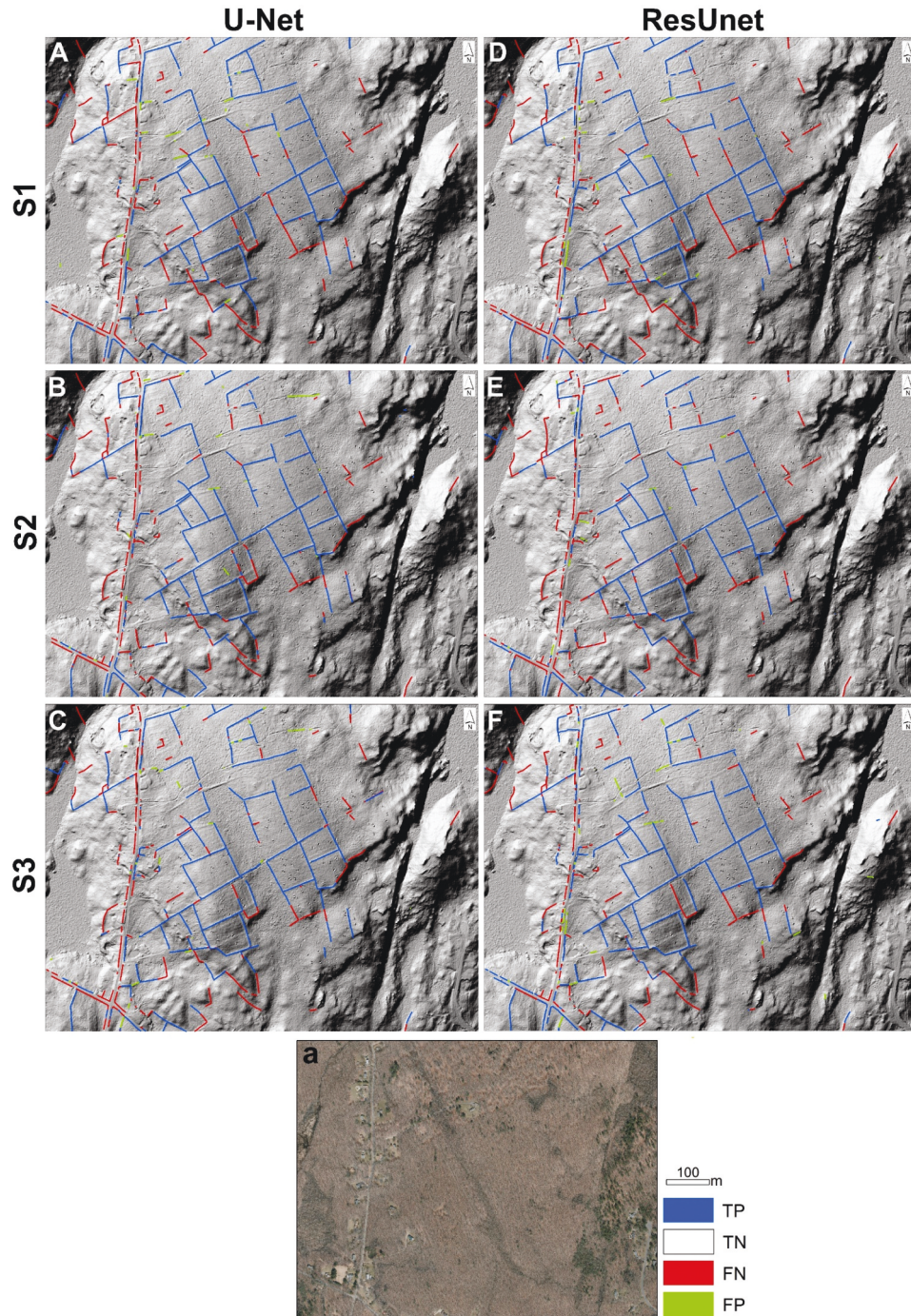
maps are the only input sources for model training and testing. In particular, we will address four factors: (1) morphological similarity, (2) sunlight angle of hillshade images, (3) shade or brightness of hillshade images, and (4) land cover and canopy. We will investigate the influence of these factors on commission and omission errors of the models in the following subsections.

#### 4.3.1. Morphological similarity

Morphological similarity to stone walls in the LiDAR image can increase the likelihood of FP (commission

error) since the model can detect non-stone walls as stone wall features. Figure 10 illustrates three examples of commission errors shown in the two test sites. Firstly, a stone wall-like feature can appear on the LiDAR hillshade images that is not an actual stone wall in the field verification (Figure 10(a,b)). This case is an inevitable error that cannot be controlled and only a field study can correct this misclassification. Secondly, linear concave features such as trail paths are another geometric shape that can confuse models (Figure 10(c,d)) because they have stone wall-like edges due to subtle slope differences



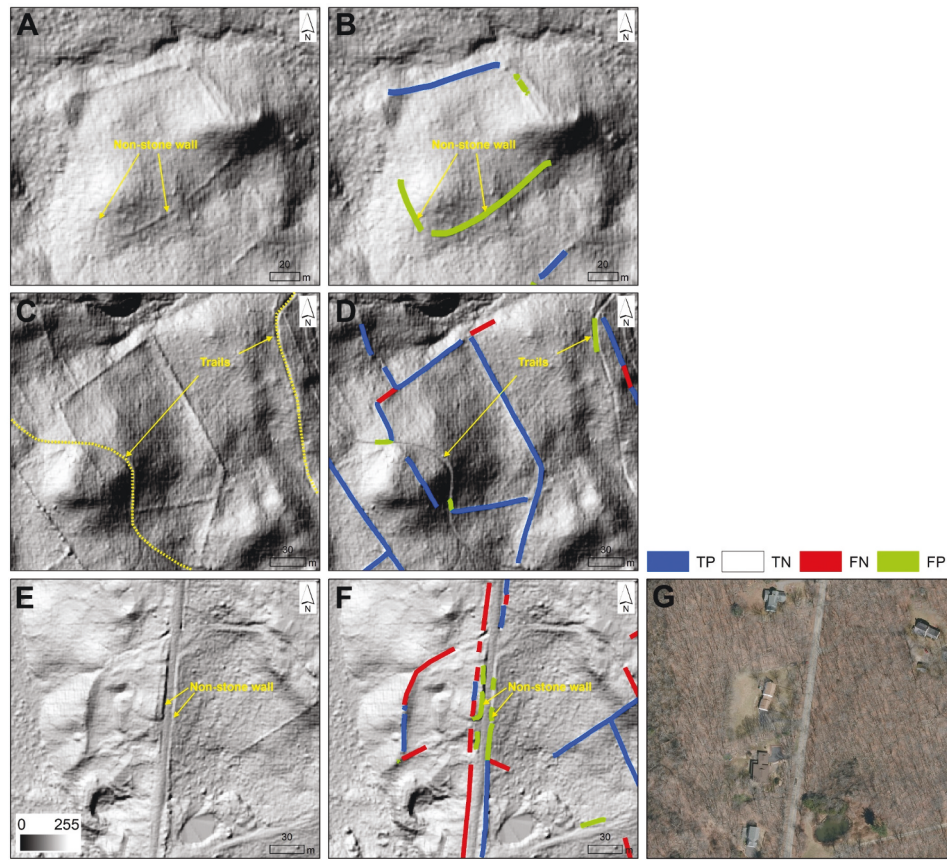


**Figure 9.** The distribution of TP, TN, FN, and FP results from different models and scenarios at test site 2 (WH). a: high-resolution orthophotography (CT ECO, 2019); A: U-Net S1; B: U-Net S2; C: U-Net S3; D: ResUnet S1; E: ResUnet S2; F: ResUnet S3. Note that TN symbol is transparent.

from the surrounding terrain. Similarly, the edges of roads and curbs also increase the chance of commission error (Figure 10(e,f)). This comprises one of the most challenging cases to accurately identify stone walls located along roadsides because these features also have height variations similar to stone walls. To mitigate the occurrence of FP resulting

from morphological similarity, one potential strategy is to leverage impervious cover data. Specifically, a buffer can be applied to exclude model-detected stone walls within certain ranges and thereby eliminate potential FP. Additional details regarding the results of this post-processing approach will be addressed in section 4.4.





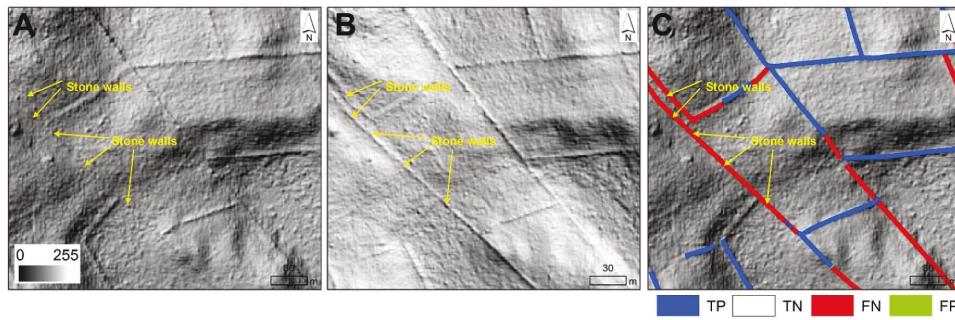
**Figure 10.** Examples of commission errors caused by morphological similarity on LiDAR hillshade surface. (A: example of wall-like features (a field edge with a pile of leaves) appearing on the hillshade map in the test site 1; B: U-Net S3 accuracy assessment result is overlaid on Figure 9(a); C: example of edge of trail in the test site 1; D: U-Net S3 accuracy assessment result is overlaid on Figure 10(c); E: example of road or curb edge in the test site 2; F: ResUnet S3 accuracy assessment result is overlaid on Figure 10(e)). G: high-resolution orthophotography (CT ECO, 2019). Note that we used the best accuracy assessment result for each test site in this figure and background image is NW hillshade map. Note that TN symbol is transparent.

#### 4.3.2. Sunlight angle of hillshade image

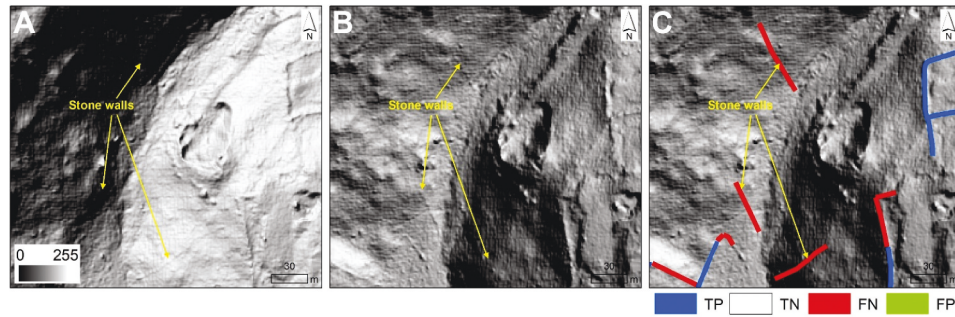
Regarding FN (omission error), the sunlight angle or azimuth angle can cause the absence of stone walls on the hillshade image. Specifically, if the direction of sunlight and stone walls is significantly well aligned, the stone wall may lose its morphological properties in the hillshade map and appear like a flat surface without shadows along the stone wall (Figure 11(a)). As described in Figure 11(a), NW-SE directional stone walls tend to be displayed as flat terrain in the NW hillshade image while they are clearly visible on the NE hillshade image, as indicated in Figure 11(b). Therefore, this factor can particularly affect the performance of model S1 since its input includes only the NW hillshade image, which shows several FNs distributed along NW-SE directional stone walls. Due to this, stacking

different directional hillshade images can help avoid missing morphological characteristics in the hillshade image, leading to accuracy improvement. The impact of this factor is supported by the result of test site 1 where it has many NW-SE directional stone walls. For example, the omission error (100% - recall score) of U-Net decreased from 29.5% (S1) to 16.5% (S3) and that of ResUnet decreased from 33% (S1) to 17.6% (S3) (Table 5).

LiDAR visualization is a crucial aspect of mapping anthropogenic features using a DCNN approach. Previous studies have shown that combining different visualization products or blending them into single raster, such as Multiscale Topographic Position and VAT can enhance model performance (Guyot, Lennon, and Hubert-Moy 2021; Kokalj and Somrak 2019; Suh et al. 2021). Therefore, further



**Figure 11.** Example of omission error for S1 in test site 1 caused by alignment for direction of sunlight and stone walls. (A: stone walls in NW-SE direction are invisible in the NW hillshade map; B: stone walls in NW-SE direction are visible in the NE hillshade map; C: U-Net S1 accuracy assessment result is overlaid on Figure 10(a)). Note that TN symbol is transparent.



**Figure 12.** Example of omission error in test site 2 caused by darkness or brightness on the hillshade map. (A: stone walls disappear due to shade or low signal in NW hillshade map; B: stone walls appear unclear due to shade or low signal in NE hillshade map; C: ResUnet S3 accuracy assessment result is overlaid on Figure 11(b)). Note that TN symbol is transparent.

research is necessary to explore how blending or combining different LiDAR visualizations can improve the identification of stone walls.

#### 4.3.3. Darkness or brightness of hillshade image

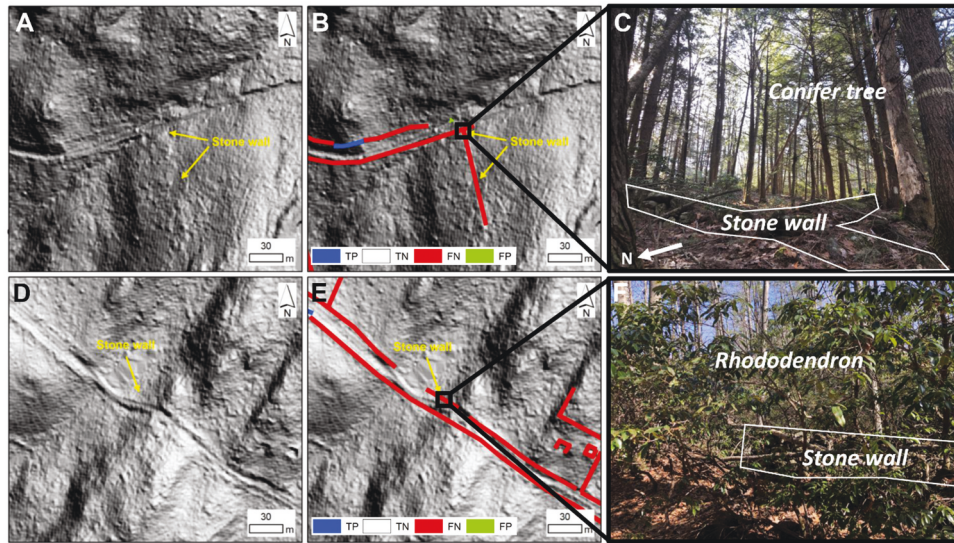
In addition to sunlight angle of hillshade image, another factor that can cause FN is a darkness or brightness of the hillshade image. For example, stone walls on shaded slopes appear as nearly black, while those on sunlight-facing slopes appear as nearly white in the hillshade map. Figure 12 illustrates FNs are distributed in extremely dark or bright areas, preventing the stone walls from appearing distinctly on the surface. Additionally, the distribution of FNs along stream in test site 1 (Figure 8) is related to the shade.

#### 4.3.4. Land cover and forest canopy

Regarding the quality of LiDAR data, land cover and forest canopy are important factors. Dense forest canopy, like that found in coniferous forests,

can restrict laser penetration and result in a sparse point spacing and lack of ground-classified points available for raster interpolation. Consequently, the quality of DEM can be poor. For test site 1 and 2, where over 80% of each site is covered by deciduous forest, the quality of LiDAR data is expected to be good since the data were collected during the leaf-off season. However, we found that the distribution of FNs near the road at test site 2 is partially associated with stone walls under coniferous trees. Coniferous trees and dense understory vegetation above stone walls prevents the laser from reaching the ground surface to collect topographic information, causing noise in the LiDAR point cloud product. To investigate the relationship between canopy cover and poor LiDAR quality, additional fieldwork was conducted at Yale Forest in Eastford town, Connecticut, which has a conifer-dominant forest. Figure 13 demonstrates the results of poor model performance, which are





**Figure 13.** Example of poor-quality LiDAR hillshade image due to stone walls under canopy cover in Eastford town, CT. (A&D: NE hillshade map; B&E: U-Net S3 accuracy assessment result; C&F: photos taken from field work). Note that TN symbol is transparent.

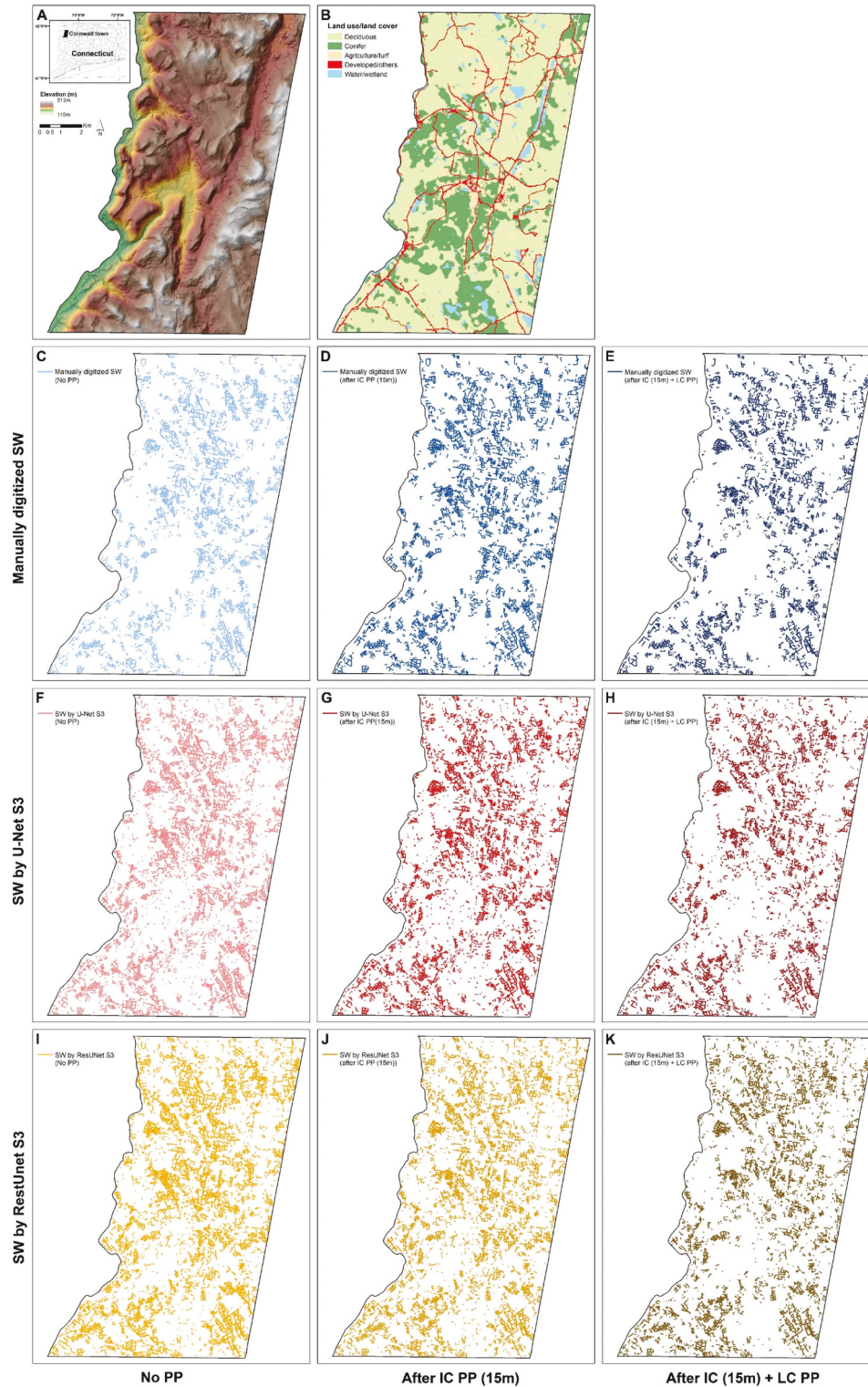
linked to the vague geometric shape of stone walls in the LiDAR hillshade due to dense canopy cover. One possible approach to address the impact of poor LiDAR quality is to utilize land cover data and exclude model results in areas that are potentially of poor quality, such as coniferous forests, wetlands, and urban areas. Section 4.4 will provide further information on the results of this post-processing strategy.

#### 4.4. Scaling up stone wall mapping

According to the results from the two test sites, the morphological similarity to stone walls and land cover type are main elements, inevitably leading to FP and FN errors. Figure 14 illustrates the stone wall map results of manual digitization, U-Net, and ResUnet model predictions, with (IC PP and IC LC PP) and without post-processing (no PP), several noisy segments distributed in coniferous forests might be related to the low quality of LiDAR derivatives. These segments are masked out after IC LC PP. Figure 15 demonstrates the accuracy assessment results for the town-level stone wall maps, showing the precision, recall, and F1 scores of U-Net S3 and ResUnet S3 for different thresholds, including the buffering of model prediction result (3 m vs. 5 m) and the buffering of

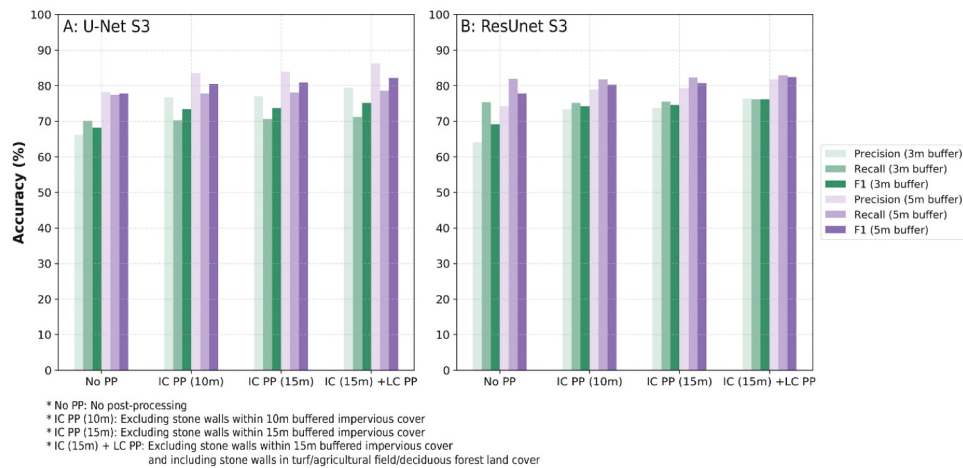
impervious cover (0 m vs. 10 m vs. 15 m). Overall, the 5 m buffering shows higher accuracy than the 3 m buffering for both model results because the larger buffering size can include more true positives with a large offset between digitized lines and model-predicted lines. Besides, excluding stone walls near impervious cover (i.e. IC PP) leads to increased accuracy, despite there being no significant difference between the 10 m and 15 m buffering thresholds. Post-processing is essential in enhancing precision by decreasing the number of FPs. This implies that FPs tend to occur in regions with impervious cover and low-quality LiDAR, which can be included in broad-scale mapping, although these examples were not well-represented in test site 1 and test site 2. IC LC PP leads to a 1.3 ~ 1.7% increase in the  $F_1$  score, indicating that post-processing significantly enhances accuracy.

Overall, the  $F_1$  score for U-Net S3 ranges from 68.1% to 82.2%, and for ResUnet S3 it ranges from 69.2% to 82.4%, depending on the thresholds and post-processing applied. Among these, IC LC PP achieved the best  $F_1$  score, which shows that our models are capable of accurately mapping high density of stone wall regions in forested areas at a broader spatial scale with over 80%  $F_1$  accuracy result. When comparing the accuracy of the U-Net and ResUnet models, there is a slight difference



**Figure 14.** Stone wall mapping result in Cornwall town, CT (A: elevation map; B: 30m resolution land use/land cover map of 2015 (data source: (Center for Land Use Education & Research CT CLEAR. 2016)); C: Manually digitized SW result without post-processing (no PP); D: Manually digitized SW result after impervious cover post-processing (IC PP); E: Manually digitized SW result after impervious cover land cover post-processing (IC LC PP); F: U-Net S3 prediction result without post-processing (no PP); G: U-Net S3 prediction result after impervious cover post-processing (IC PP); H: U-Net S3 prediction result after impervious cover land cover post-processing (IC LC PP); I: ResUnet S3 prediction result without post-processing (no PP); J: ResUnet S3 prediction result after IC PP; K: ResUnet S3 prediction result after IC LC PP).





**Figure 15.** Accuracy assessment results of U-Net S3 and ResUnet S3 for town-level stone wall mapping.

(0.2%) in the best  $F_1$  score (IC LC PP). Regarding the accuracy trend, the U-Net model tends to show high precision and relatively low recall score, while ResUnet tends to show high recall and relatively low precision. This implies that U-Net model tends to be conservative in mapping stone walls, while ResUnet has a better balance between precision and recall.

## 5. Conclusions

This study presents an innovative framework that uses U-Net-based models to automatically detect linear anthropogenic landscape features, particularly stone walls, from high-resolution airborne LiDAR derivatives. Our study draws three main conclusions. First, the DCNN models play a key role in recording anthropogenic landforms and mapping the extent of historic land use legacy. Our models identified stone walls in forested areas with a best  $F_1$  score of 88.3% (U-Net S3) at test site 1 and  $F_1$  score of 80.3% (ResUnet S3) at test site 2. Regarding town-level mapping, both models performed similarly to the two local scale forested areas (test 1 and test 2) with a  $F_1$  score over 82% after post-processing. Second, we found four factors that need to be considered when implementing LiDAR imagery to detect small-scale geomorphic features in DCNN models: (1) morphological similarity to the target feature, (2) sunlight angle of the hillshade image, (3) darkness or brightness of the hillshade image, and (4) land cover and forest canopy. In particular, land cover can directly affect light or laser penetration of

LiDAR sensor, which results in the quality of LiDAR image and model performance. Therefore, our trained models are suitable for detecting stone walls in reforested area with deciduous dominant forests that had used for agriculture or pasture in 17th to early 20th centuries. Third, our proposed method will increase stone wall detection and rapidly expand the scale of investigation of anthropogenic features in the Northeastern USA, which has important implications in studying anthropogenic impacts on geomorphology, forest structure, and ecology.

## Notes

1. CT stone wall mapper: <https://connecticut.maps.arcgis.com/apps/webappviewer/index.html?id=0208461aa98a4df3969624e7110e1f2c>.
2. NH stone wall mapper: <https://nhdes.maps.arcgis.com/apps/webappviewer/index.html?id=f4d57ec1a6b8414190ca0662456dfffb0>.

## Acknowledgments

The authors wish to thank Jonathan Leonard for field mapping and Katharine Johnson for comments on this contribution. The authors would like to thank three anonymous reviewers for their valuable feedback and comments that greatly improved the quality of this contribution.

## Disclosure statement

No potential conflict of interest was reported by the author(s).



## Funding

This work was funded by National Science Foundation grant BCS-1654462 to W. Ouimet.

## ORCID

Ji Won Suh  <http://orcid.org/0000-0002-6406-0024>

## Data availability statement

LiDAR data that used for this study are available upon request by contact with the corresponding author, or accessed through <https://cteco.uconn.edu/data/download/flight2016/index.htm>.

## References

- Abdollahi, A., B. Pradhan, N. Shukla, S. Chakraborty, and A. Alamri. 2020. "Deep Learning Approaches Applied to Remote Sensing Datasets for Road Extraction: A State-Of-The-Art Review." *Remote Sensing* 12 (9): 1444. doi:10.3390/RS12091444.
- Bennett, R., K. Welham, R. A. Hill, and A. Ford. 2012. "A Comparison of Visualization Techniques for Models Created from Airborne Laser Scanned Data." *Archaeological Prospection* 19 (1): 41–48. doi:10.1002/arp.1414.
- Blaschke, T., S. Lang, E. Lorup, J. Strobl, and P. Zeil. 2000. "Object-Oriented Image Processing in an Integrated GIS/Remote Sensing Environment and Perspectives for Environmental Applications. Environmental Information for Planning, Politics and the Public." *Environmental Information for Planning, Politics and the Public* 2 (1995): 555–570.
- Capitol Region Council of Governments (CRCoG). 2016a. "2016 Aerial Imagery." <http://Cteco.Uconn.Edu/Data/Flight2016/Index.Htm>
- Capitol Region Council of Governments (CRCoG). 2016b. "Connecticut Statewide LiDAR 2016 Bare Earth DEM." [http://www.Cteco.Uconn.Edu/Metadata/Dep/Document/LidarDEM\\_2016\\_fgdc\\_plus.Htm](http://www.Cteco.Uconn.Edu/Metadata/Dep/Document/LidarDEM_2016_fgdc_plus.Htm)
- Carbonneau, P. E., S. J. Dugdale, T. P. Breckon, J. T. Dietrich, M. A. Fonstad, H. Miyamoto, and A. S. Woodget. 2020. "Adopting Deep Learning Methods for Airborne RGB Fluvial Scene Classification." *Remote Sensing of Environment* 251 (September): 112107. Elsevier. doi:10.1016/j.rse.2020.112107.
- Chen, Y., G. He, R. Yin, K. Zheng, and G. Wang. 2022. "Comparative Study of Marine Ranching Recognition in Multi-Temporal High-Resolution Remote Sensing Images Based on DeepLab-V3+ and U-Net." *Remote Sensing* 14 (22): 5654. MDPI. doi:10.3390/rs14225654.
- Chen, L. C., G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. 2018. "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40 (4): 834–848. doi:10.1109/TPAMI.2017.2699184.
- Chen, Y., Y. Wei, Q. Wang, F. Chen, C. Lu, and S. Lei. 2020. "Mapping Post-Earthquake Landslide Susceptibility: A U-Net Like Approach." *Remote Sensing* 12 (17): 2767. doi:10.3390/RS12172767.
- Chicco, D., and G. Jurman. 2020. "The Advantages of the Matthews Correlation Coefficient (MCC) Over F1 Score and Accuracy in Binary Classification Evaluation." *BMC Genomics* 21 (1). BioMed Central. doi:10.1186/s12864-019-6413-7.
- Connecticut Center for Land Use Education & Research (CT CLEAR). 2016. "2015 Connecticut Land Cover." <https://Clear.Uconn.Edu/Projects/Landscape/Download.Htm#top>
- Connecticut Center for Land Use Education & Research (CT CLEAR). 2017. "Statewide Impervious 2012."
- Connecticut Environmental Conditions Online (CT ECO). 2016. "Connecticut Statewide LiDAR 2016."
- Connecticut Environmental Conditions Online (CT ECO). 2019. "Connecticut Orthoimagery 2019."
- Crow, P., S. Benham, B. J. Devereux, and G. S. Amable. 2007. "Woodland Vegetation and Its Implications for Archaeological Survey Using LiDar." *Forestry* 80 (3): 241–252. doi:10.1093/forestry/cpm018.
- Davis, D. S., C. P. Lipo, and M. C. Sanger. 2019. "A Comparison of Automated Object Extraction Methods for Mound and Shell-Ring Identification in Coastal South Carolina." *Journal of Archaeological Science: Reports* 23 (November 2018): 166–177. Elsevier. doi:10.1016/j.jasrep.2018.10.035.
- Diakogiannis, F. I., F. Waldner, P. Caccetta, and C. Wu. 2020. "ResUnet-A: A Deep Learning Framework for Semantic Segmentation of Remotely Sensed Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 162 (January): 94–114. Elsevier. doi:10.1016/j.isprsjprs.2020.01.013.
- Doneus, M. 2013. "Openness as Visualization Technique for Interpretative Mapping of Airborne Lidar Derived Digital Terrain Models." *Remote Sensing* 5 (12): 6427–6442. doi:10.3390/rs5126427.
- Doneus, M., C. Briese, M. Fera, and M. Janner. 2008. "Archaeological Prospection of Forested Areas Using Full-Waveform Airborne Laser Scanning." *Journal of Archaeological Science* 35 (4): 882–893. doi:10.1016/j.jas.2007.06.013.
- Evans, D. H., R. J. Fletcher, C. Pottier, J. B. Chevance, D. Soutif, B. S. Tan, S. Im, et al. 2013. "Uncovering Archaeological Landscapes at Angkor Using Lidar." *Proceedings of the National Academy of Sciences of the United States of America* 110 (31): 12595–12600. doi:10.1073/pnas.1306539110.
- Foster, D. R., B. Donahue, D. Kittredge, G. Motzkin, B. Hall, B. Turner, and E. Chilton. 2008. "New England's Forest Landscape: Ecological Legacies and Conservation Patterns Shaped by Agrarian History." In *Agrarian Landscapes in Transition: Comparisons of Long-Term Ecological and Cultural Change*, 44–88. Oxford University Press.
- Freeland, T., B. Heung, D. V. Burley, G. Clark, and A. Knudby. 2016. "Automated Feature Extraction for Prospection and Analysis of Monumental Earthworks from Aerial LiDar in the

- Kingdom of Tonga." *Journal of Archaeological Science* 69: 64–74. Elsevier. doi:10.1016/j.jas.2016.04.011.
- Gallwey, J., M. Eyre, M. Tonkins, and J. Coggan. 2019. "Bringing Lunar LiDar Back Down to Earth: Mapping Our Industrial Heritage Through Deep Transfer Learning." *Remote Sensing* 11 (17): 1994. doi:10.3390/rs11171994.
- Girshick, R., J. Donahue, T. Darrell, and J. Malik. 2013. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." November. <http://arxiv.org/abs/1311.2524>
- Guyot, A., M. Lennon, and L. Hubert-Moy. 2021. "Combined Detection and Segmentation of Archeological Structures from LiDar Data Using a Deep Learning Approach." *Journal of Computer Applications in Archaeology* 4 (1): 1–19. doi:10.5334/jcaa.64.
- He, K., G. Gkioxari, P. Dollár, and R. Girshick. 2020. "Mask R-CNN." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42 (2): 386–397. doi:10.1109/TPAMI.2018.2844175.
- Heidemann, H. 2018. *Lidar Base Specification*.
- Hesse, R. 2010. "LiDAR-Derived Local Relief Models-A New Tool for Archaeological Prospection." *Archaeological Prospection* 17 (2): 67–72. doi:10.1002/arp.374.
- He, K., X. Zhang, S. Ren, and J. Sun. 2015. "Deep Residual Learning for Image Recognition." In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. doi:10.1109/CVPR.2016.90.
- Howey, M. C. L., F. B. Sullivan, J. Tallant, R. V. Kopple, and M. W. Palace. 2016. "Detecting Precontact Anthropogenic Microtopographic Features in a Forested Landscape with Lidar: A Case Study from the Upper Great Lakes Region, AD 1000-1600." *Plos One* 11 (9): 1–11. doi:10.1371/journal.pone.0162062.
- Ioffe, S., and C. Szegedy. 2015. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift." *32nd International Conference on Machine Learning, ICML 2015* 1: 448–456.
- Johnson, K., and W. Ouimet. 2014. "Rediscovering the Lost Archaeological Landscape of Southern New England Using Airborne Light Detection and Ranging (LiDar)." *Journal of Archaeological Science* 43 (1): 9–20. Elsevier. doi:10.1016/j.jas.2013.12.004.
- Johnson, K., and W. Ouimet. 2016. "Physical Properties and Spatial Controls of Stone Walls in the Northeastern USA: Implications for Anthropocene Studies of 17th to Early 20th Century Agriculture." *Anthropocene* 15: 22–36. Elsevier B.V. doi:10.1016/j.ancene.2016.07.001.
- Johnson, K., and W. Ouimet. 2018. "An Observational and Theoretical Framework for Interpreting the Landscape Palimpsest Through Airborne LiDar." *Applied Geography* 91 (January): 32–44. Elsevier. doi:10.1016/j.apgeog.2017.12.018.
- Johnson, K., and W. Ouimet. 2021. "Reconstructing Historical Forest Cover and Land Use Dynamics in the Northeastern United States Using Geospatial Analysis and Airborne LiDar." *Annals of the American Association of Geographers* 111 (6): 1656–1678. Taylor & Francis. doi:10.1080/24694452.2020.1856640.
- Johnson, K., W. Ouimet, S. Dow, and C. Haverfield. 2021. "Estimating Historically Cleared and Deforested Land in Massachusetts Using Airborne LiDar." *Remote Sensing* 13 (21): 4318. doi:10.3390/rs13214318.
- Kim, M., J. Lee, D. Han, M. Shin, J. Im, J. Lee, L. J. Quackenbush, and Z. Gu. 2018. "Convolutional Neural Network-Based Land Cover Classification Using 2-D Spectral Reflectance Curve Graphs with Multitemporal Satellite Imagery." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11 (12): 4604–4617. IEEE. doi:10.1109/JSTARS.2018.2880783.
- Kingma, D. P., and J. Ba. 2015. "Adam: A Method for Stochastic Optimization." ICLR.
- Kokalj, Ž., and M. Somrak. 2019. "Why Not a Single Image? Combining Visualizations to Facilitate Fieldwork and On-Screen Mapping." *Remote Sensing* 11 (7): 747. doi:10.3390/rs11070747.
- Kokalj, Ž., K. Zaksek, and K. Oštir. 2013. "Visualizations of Lidar Derived Relief Models." In *Interpreting Archaeological Topography: Airborne Laser Scanning, 3D Data and Ground Observation*, edited by R. S. Opitz and D. C. Cowley, 100–114. Oxford, England: Oxbow Books.
- Kugelman, J., J. Allman, S. A. Read, S. J. Vincent, J. Tong, M. Kalloniatis, F. K. Chen, M. J. Collins, and D. Alonso-Caneiro. 2022. "A Comparison of Deep Learning U-Net Architectures for Posterior Segment OCT Retinal Layer Segmentation." *Scientific Reports* 12(1). Nature Research. doi:10.1038/s41598-022-18646-2.
- Kussul, N., M. Lavreniuk, S. Skakun, and A. Shelestov. 2017. "Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data." *IEEE Geoscience and Remote Sensing Letters* 14 (5): 778–782. IEEE. doi:10.1109/LGRS.2017.2681128.
- Leonard, J., W. B. Ouimet, and S. Dow. 2021. "Evaluating User Interpretation and Error Associated with Digitizing Stone Walls Using Airborne LiDar." *Geological Society of America Abstracts with Programs* 53 (1). doi:10.1130/abs/2021NE-361715.
- Li, W., R. Buitenwerf, M. Munk, P. K. Bøcher, and J. C. Svenning. 2020. "Deep-Learning Based High-Resolution Mapping Shows Woody Vegetation Densification in Greater Maasai Mara Ecosystem." *Remote Sensing of Environment* 247 (November 2019): 111953. Elsevier. doi:10.1016/j.rse.2020.111953.
- Liu, P., Y. Wei, Q. Wang, Y. Chen, and J. Xie. 2020. "Research on Post-Earthquake Landslide Extraction Algorithm Based on Improved U-Net Model." *Remote Sensing* 12 (5): 1–14. doi:10.3390/rs12050894.
- Magnini, L., C. Bettineschi, and A. De Guio. 2017. "Object-Based Shell Craters Classification from LiDAR-Derived Sky-View Factor." *Archaeological Prospection* 24 (3): 211–223. doi:10.1002/arp.1565.
- Mboga, N., T. Grippa, S. Georganos, S. Vanhuyse, B. Smets, O. Dewitte, E. Wolff, and M. Lennert. 2020. "Fully Convolutional Networks for Land Cover Classification from

- Historical Panchromatic Aerial Photographs." *ISPRS Journal of Photogrammetry and Remote Sensing* 167 (July): 385–395. Elsevier. doi:10.1016/j.isprsjprs.2020.07.005.
- Niculiță, M. 2020. "Geomorphometric Methods for Burial Mound Recognition and Extraction from High-Resolution LiDAR DEMs." *Sensors (Switzerland)* 20 (4): 1192. doi:10.3390/s20041192.
- Olteanu-Raimond, A. M., L. See, M. Schultz, G. Foody, M. Riffler, T. Gasber, L. Jolivet, et al. 2020. "Use of Automated Change Detection and VGI Sources for Identifying and Validating Urban Land Use Change." *Remote Sensing* 12 (7): 1186. doi:10.3390/rs12071186.
- Orongo, H. A., F. C. Conesa, A. Garcia-Molsosa, A. Lobo, A. S. Green, M. Madella, and C. A. Petrie. 2020. "Automated Detection of Archaeological Mounds Using Machine-Learning Classification of Multisensor and Multitemporal Satellite Data." *Proceedings of the National Academy of Sciences of the United States of America* 117 (31): 18240–18250. doi:10.1073/pnas.2005583117.
- Qi, W., M. Wei, W. Yang, C. Xu, and C. Ma. 2020. "Automatic Mapping of Landslides by the ResU-Net." *Remote Sensing* 12 (15): 1–15. doi:10.3390/RS12152487.
- Ren, S., K. He, R. Girshick, and J. Sun. 2017. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39 (6): 1137–1149. IEEE. doi:10.1109/TPAMI.2016.2577031.
- Ridge, J. C. 2004. "The Quaternary glaciation of western New England with correlations to surrounding areas." *Developments in Quaternary Sciences*, 2: 169–199. doi:10.1016/S1571-0866(04)80196-9.
- Risbøl, O., and L. Gustavsen. 2018. "LiDAR from Drones Employed for Mapping Archaeology – Potential, Benefits and Challenges." *Archaeological Prospection* 25 (4): 329–338. doi:10.1002/arp.1712.
- Rodríguez, C. G., J. Vitrià, and O. Mora. 2020. "Uncertainty-Based Human-In-The-Loop Deep Learning for Land Cover Segmentation." *Remote Sensing* 12 (22): 1–14. doi:10.3390/rs12223836.
- Ronneberger, O., P. Fischer, and T. Brox. 2015. "U-Net: Convolutional Networks for Biomedical Image Segmentation." *Medical Image Computing and Computer-Assisted Intervention* 9351: 234–241. doi:10.1007/978-3-319-24574-4\_28.
- Schneider, A., M. Takla, A. Nicolay, A. Raab, and T. Raab. 2015. "A Template-Matching Approach Combining Morphometric Variables for Automated Mapping of Charcoal Kiln Sites." *Archaeological Prospection* 22 (1): 45–62. doi:10.1002/arp.1497.
- Srivastava, S., J. E. Vargas-Muñoz, and D. Tuia. 2019. "Understanding Urban Landuse from the Above and Ground Perspectives: A Deep Learning, Multimodal Solution." *Remote Sensing of Environment* 228 (October 2018): 129–143. Elsevier. doi:10.1016/j.rse.2019.04.014.
- Stoian, A., V. Poulain, J. Inglada, V. Poughon, and D. Derksen. 2019. "Land Cover Maps Production with High Resolution Satellite Image Time Series and Convolutional Neural Networks: Adaptations and Limits for Operational Systems." *Remote Sensing* 11 (17): 1–26. doi:10.3390/rs11171986.
- Štular, B., Ž. Kokalj, K. Oštir, and L. Nuninger. 2012. "Visualization of Lidar-Derived Relief Models for Detection of Archaeological Features." *Journal of Archaeological Science* 39 (11): 3354–3360. doi:10.1016/j.jas.2012.05.029.
- Suh, J. W., E. Anderson, W. Ouimet, K. Johnson, and C. Witharana. 2021. "Mapping Relict Charcoal Hearths in the Northeast US Using Deep Learning Convolutional Neural Networks and LIDAR Data." *Remote Sensing* 13 (22): 4630. doi:10.3390/rs13224630.
- Sylvain, J. D., G. Drolet, and N. Brown. 2019. "Mapping Dead Forest Cover Using a Deep Convolutional Neural Network and Digital Aerial Photography." *ISPRS Journal of Photogrammetry and Remote Sensing* 156 (August): 14–26. Elsevier. doi:10.1016/j.isprsjprs.2019.07.010.
- Tan, Y., S. Xiong, and Y. Li. 2018. "Automatic Extraction of Built-Up Areas from Panchromatic and Multispectral Remote Sensing Images Using Double-Stream Deep Convolutional Neural Networks." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11 (11): 3988–4004. IEEE. doi:10.1109/JSTARS.2018.2871046.
- Thorson, R. M. 2002. *Stone by Stone : The Magnificent History In New England's Stone Walls*. New York: Walker & Company.
- Thorson, R. M. 2023. "Taxonomy and Nomenclature of the Stone Domain in New England." *Historical Archaeology*.
- Trier, Ø. D., J. H. Reksten, and K. Løseth. 2021. "Automated Mapping of Cultural Heritage in Norway from Airborne Lidar Data Using Faster R-CNN." *International Journal of Applied Earth Observation and Geoinformation* 95: 102241. doi:10.1016/j.jag.2020.102241.
- Trier, Ø. D., M. Zortea, and C. Tønning. 2015. "Automatic Detection of Mound Structures in Airborne Laser Scanning Data." *Journal of Archaeological Science: Reports* 2 (February): 69–79. doi:10.1016/j.jasrep.2015.01.005.
- Verbovšek, T., T. Popit, and Ž. Kokalj. 2019. "VAT Method for Visualization of Mass Movement Features: An Alternative to Hillshaded DEM." *Remote Sensing* 11 (24): 1–14. doi:10.3390/rs11242946.
- Verschoof van der Vaart, W., A. Bonhage, A. Schneider, W. Ouimet, and T. Raab. 2022. "Automated Large-Scale Mapping and Analysis of Relict Charcoal Hearths in Connecticut (USA) Using a Deep Learning YOLOv4 Framework." *Archaeological Prospection* John Wiley and Sons. doi:10.1002/arp.1889.
- Verschoof-van der Vaart, W., and K. Lambers. 2019. "Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands." *Journal of Computer Applications in Archaeology* 2 (1): 31–40. doi:10.5334/jcaa.32.
- Verschoof-Van Der Vaart, W. B., K. Lambers, W. Kowalczyk, and Q. P. J. Bourgeois. 2020. "Combining Deep Learning and Location-Based Ranking for Large-Scale Archaeological Prospection of LiDAR Data from the Netherlands." *ISPRS International Journal of Geo-Information* 9 (5): 293. doi:10.3390/ijgi9050293.



- Verschoof-van der Vaart, W., and J. Landauer. 2021. "Using CarcassNet to Automatically Detect and Trace Hollow Roads in LiDAR Data from the Netherlands." *Journal of Cultural Heritage* 47 (January): 143–154. Elsevier Masson s.r. l. doi:[10.1016/j.culher.2020.10.009](https://doi.org/10.1016/j.culher.2020.10.009).
- Waldner, F., and F. I. Diakogiannis. 2020. "Deep Learning on Edge: Extracting Field Boundaries from Satellite Images with a Convolutional Neural Network." *Remote Sensing of Environment* 245 (October 2019): 111741. Elsevier. doi:[10.1016/j.rse.2020.111741](https://doi.org/10.1016/j.rse.2020.111741).
- Wang, Y., and M. Li. 2019. "Urban Impervious Surface Detection from Remote Sensing Images: A Review of the Methods and Challenges." *IEEE Geoscience and Remote Sensing Magazine* 7 (3): 64–93. IEEE. doi:[10.1109/MGRS.2019.2927260](https://doi.org/10.1109/MGRS.2019.2927260).
- Wessels, T. 1997. *Reading the Forested Landscape*. New York: The Countryman Press.
- Witharana, C., W. Ouimet, and K. Johnson. 2018. "Using LiDAR and GEOBIA for Automated Extraction of Eighteenth–Late Nineteenth Century Relict Charcoal Hearths in Southern New England." *GIScience & Remote Sensing* 55 (2): 183–204. Taylor & Francis. doi:[10.1080/15481603.2018.1431356](https://doi.org/10.1080/15481603.2018.1431356).
- Wu, R., G. Liu, R. Zhang, X. Wang, Y. Li, and B. Zhang. 2020. "A Deep Learning Method for Mapping Glacial Lakes from the Combined Use of Synthetic-Aperture Radar and Optical Satellite Images." *Remote Sensing* 12 (24): 4020. DI. doi:[10.3390/rs12244020](https://doi.org/10.3390/rs12244020).
- Yan, S., L. Xu, G. Yu, L. Yang, W. Yun, D. Zhu, S. Ye, and X. Yao. 2021. "Glacier Classification from Sentinel-2 Imagery Using Spatial-Spectral Attention Convolutional Model." *International Journal of Applied Earth Observation and Geoinformation* 102: 102445. Elsevier B.V. doi:[10.1016/j.jag.2021.102445](https://doi.org/10.1016/j.jag.2021.102445).
- Zhang, C., P. Harrison, X. Pan, H. Li, I. Sargent, and P. Atkinson. 2020. "Scale Sequence Joint Deep Learning (SS-JDL) for Land Use and Land Cover Classification." *Remote Sensing of Environment* 237 (September 2019): 111593. Elsevier. doi:[10.1016/j.rse.2019.111593](https://doi.org/10.1016/j.rse.2019.111593).
- Zhang, Z., Q. Liu, and Y. Wang. 2018. "Road Extraction by Deep Residual U-Net." *IEEE Geoscience and Remote Sensing Letters* 15 (5): 749–753. doi:[10.1109/LGRS.2018.2802944](https://doi.org/10.1109/LGRS.2018.2802944).
- Zhang, C., I. Sargent, X. Pan, H. Li, A. Gardiner, J. Hare, and P. M. Atkinson. 2018. "An Object-Based Convolutional Neural Network (OCNN) for Urban Land Use Classification." *Remote Sensing of Environment* 216 (January): 57–70. Elsevier. doi:[10.1016/j.rse.2018.06.034](https://doi.org/10.1016/j.rse.2018.06.034).
- Zhong, L., L. Hu, and H. Zhou. 2019. "Deep Learning Based Multi-Temporal Crop Classification." *Remote Sensing of Environment* 221 (November 2018): 430–443. Elsevier. doi:[10.1016/j.rse.2018.11.032](https://doi.org/10.1016/j.rse.2018.11.032).
- Zhou, Z., M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang. 2018. "Unet++: A Nested U-Net Architecture for Medical Image Segmentation." *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 11045 LNCS. Springer International Publishing. doi:[10.1007/978-3-030-00889-5\\_1](https://doi.org/10.1007/978-3-030-00889-5_1).