

2



Article Semantic Segmentation (U-Net) of Archaeological Features in Airborne Laser Scanning—Example of the Białowieża Forest

Paweł Zbigniew Banasiak ^{1,*,†}, Piotr Leszek Berezowski ^{1,†}, Rafał Zapłata ², Miłosz Mielcarek ³, Konrad Duraj ⁴ and Krzysztof Stereńczak ³

- ¹ Data Processing Lab, 40748 Katowice, Poland; piotr@dataprocessinglab.com
- School of Exact Sciences, Faculty of Mathematics and Natural Sciences, Cardinal Wyszyński University in Warsaw, 01938 Warszawa, Poland; r.zaplata@uksw.edu.pl
- ³ Department of Geomatics, Forest Research Institute, 05090 Sekocin Stary, Poland; m.mielcarek@ibles.waw.pl (M.M.); k.sterenczak@ibles.waw.pl (K.S.)
- ⁴ Department of Biosensors and Processing of Biomedical Signals, Faculty of Biomedical Engineering, Silesian University of Technology, 41800 Zabrze, Poland; konrad.duraj@polsl.pl
- * Correspondence: pawel@dataprocessinglab.com
- † These authors contributed equally to this work.

Abstract: Airborne Laser Scanning (ALS) technology can be used to identify features of terrain relief in forested areas, possibly leading to the discovery of previously unknown archaeological monuments. Spatial interpretation of numerous objects with various shapes and sizes is a difficult challenge for archaeologists. Mapping structures with multiple elements whose area can exceed dozens of hectares, such as ancient agricultural field systems, is very time-consuming. These archaeological sites are composed of a large number of embanked fields, which together form a recognizable spatial pattern. Image classification and segmentation, as well as object recognition, are the most important tasks for deep learning neural networks (DLNN) and therefore they can be used for automatic recognition of archaeological monuments. In this study, a U-Net neural network was implemented to perform semantic segmentation of the ALS-derived data including (1) archaeological, (2) natural and (3) modern features in the Polish part of the Białowieża Forest. The performance of the U-Net segmentation model was evaluated by measuring the pixel-wise similarity between ground truth and predicted segmentation masks. After 83 epochs, The Dice-Sorensen coefficient (F1 score) and the Intersect Over Union (IoU) metrics were 0.58 and 0.5, respectively. The IoU metric reached a value of 0.41, 0.62 and 0.62 for the ancient field system banks, ancient field system plots and burial mounds, respectively. The results of the U-Net deep learning model proved very useful in semantic segmentation of images derived from ALS data.

Keywords: deep neural networks; convolutional neural network; U-Net; ancient field systems; Celtic fields; Polish part of the Białowieża Forest; UNESCO World Heritage Site; LiDAR; ALS; DTM

1. Introduction

The archaeological heritage of forested areas representing traces of past human activities remains poorly researched and inventoried. Conducting an archaeological inventory is limited by available research methods. In recent years, Airbone Laser Scanning (ALS) technology has greatly increased the chance for discovering archaeological objects in these areas. In Poland, forests cover over 30% of the country's territory and the ALS data of the ISOK system (Informatyczny System Osłony Kraju) cover the entire national territory.

The most common method of detecting archaeological objects in forested areas is their visual recognition and identification based on ALS data processing. Techniques based on manual labelling and georeferencing of objects require a lot of desk work and are very time-consuming. Therefore, the process of ALS data analysis has been supported by human-supervised procedures for semi-automatic or automatic recognition of potential



Citation: Banasiak, P.Z.; Berezowski, P.L.; Zapłata, R.; Mielcarek, M.; Duraj, K.; Stereńczak, K. Semantic Segmentation (U-Net) of Archaeological Features in Airborne Laser Scanning—Example of the Białowieża Forest. *Remote Sens.* 2022, 14, 995. https://doi.org/10.3390/ rs14040995

Academic Editors: Antonio M. Rinaldi and Beniamino Murgante

Received: 1 January 2022 Accepted: 14 February 2022 Published: 18 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). archaeological objects [1–3]. Modern systems for image analysis and classification are based on the so-called Deep Neural Networks (DNN). The most important feature of deep neural networks is their ability to learn from examples and the possibility of automatic generalisation of the acquired knowledge [4]. The discussed methods fall into the category of non-destructive activities, the application of which is expected mainly in those areas where invasive methods should not or cannot be used, e.g., intact cultural heritage sites, nature reserves or difficult access sites.

In order to develop automatic recognition of archaeological objects, deep learning image segmentation methods were applied to the Polish part of the Białowieża Forest (UNESCO World Heritage Site). This area has been covered by an ALS survey acquired during two photogrammetric missions—as part of the Life+ ForBioSensing project and as part of the ISOK system [5,6]. Extensive archaeological research has been carried out in recent years, mainly using the ALS data [7,8].

Due to the availability of ALS data covering the entire territory of Poland, it is now possible to conduct surveys over large areas. This prompts the development of a methodology for large-scale activities with a standardized formula for data processing and the use of neural networks to identify archaeological monuments and monitor their condition.

Therefore, the aim of the research was to implement deep neural network for automatic object recognition from preprocessed ALS data for large areas. The specific goals were: (1) to develop a system based on deep neural networks to automate the process of recognition of multiple classes of objects (archaeological, natural and modern); and (2) to evaluate the performance of deep neural networks in the recognition of archaeological monuments.

1.1. Deep Learning

Deep learning (DL) is a subcategory of Machine Learning (ML) and, more broadly, of Artificial Intelligence (AI). Considering archaeological heritage, DL offers new possibilities of archaeological monuments recognition. Large amounts of data can be analysed in a short time, thus improving the manual work of specialists and enriching the interpretation process.

Deep learning methods have dominated computer vision in recent years. Deep learning is based on multi-layer computing systems that are capable of learning and recognizing data at multiple levels of abstraction [9]. The main mathematical operation used in DNNs is convolution. Since the introduction of (CNNs) by LeCun in the early 1990s, these networks have reached a high level of performance [10]. CNNs outperform all other techniques in image recognition. A convolutional neural network compresses input images into a vector representation, which can then be used for classification, recognition or segmentation processes. The vectors, called feature vectors, represent the objects in the analysed image. Features extracted from the successive layers are combined hierarchically to recognize higher order functions [11,12]. A typical CNN consists of several layers, which can be divided into the input, the hidden and the output layers [13].

The most important advance in CNNs came after Krizhevsky created 'AlexNet', a CNN model that won the ILSVRC (ImageNet Large Scale Visual Recognition Challenge) in 2012. AlexNet classified 1.2 million images into the 1000 classes with an error rate of 15.3% [13]. This is one of the most influential papers in the field of deep learning and image recognition. The rectified linear unit (ReLU) and dropout layer techniques became standard features widely used in deep learning networks. The implementation of 2D (two-dimensional) convolution in the graphics processing unit (GPU) and the increase of available memory by using two coupled graphics cards yields impressive results.

Recent CNNs developments mainly stem from increasing the number of network layers to improve the learning depth, a crucial factor for difficult non-trivial data [14,15]. Deep neural networks are much more difficult to train since they overfit the training data and perform poorly on test data and also because gradients can disappear or explode. A significant improvement with respect to the above-mentioned issue was achieved in ResNet

(Residual Neural Network) by using so-called skip connections. ResNet won the ILSVRC competition in 2015 with an error rate of 3.57% [16].

Deep learning networks in computer vision generally perform image classification, object recognition and image segmentation. These methods help estimate the probability that an object is present in the analysed image, indicate its location with a 'bounding box' and obtain an outline called a 'mask'. Method selection depends on the requirements of the project and requires adequate preparation of the input data. Semantic image segmentation involves assigning a predefined classification label to each pixel.

In the last decade, image segmentation methods using deep machine learning solutions have evolved significantly. The first approach to such solutions was a study of so-called 'Fully Convolutional Networks' (FCN) [17]. Such networks only contain convolutional layers and are therefore capable of processing images of arbitrary size and generating a segmentation mask. By using so-called 'skip connections', it is possible to combine information from deeper layers of the network (semantic information) with features extracted in the initial layers. This architecture pioneered the use of deep learning networks for image segmentation problems. Although the model produced state-of-the-art results when used on available datasets ('benchmarks'), it suffered from problems such as long inference times and the inability to process data in so-called 'real time'. In addition, the solution did not take the global context into account and was difficult to implement with 3D data. The next step in the exploration of semantic segmentation methods that addressed the problems of FCN was an encoder-decoder network called SegNet [18]. In this solution, the segmentation mask is created step by step through successive layers of deconvolution. The U-Net created by Ronneberger, Fischer and Bronx in 2015 [19] was the next big step in image segmentation with neural networks. Similar to SegNet, the U-Net architecture consists of two parts, namely an encoder and a decoder. However, when processing the feature vector through successive deconvolution layers, the features generated by the encoder and decoder are combined to avoid the loss of spatial information. The authors proposed a training strategy of data augmentation to train the model to more efficiently extract an image feature from a small number of samples.

1.2. Deep Learning and ALS Data in Archaeology

The interest in DCNNs (Deep Convolutional Neural Networks) for the recognition of archaeological structures is gradually increasing among the archaeological community. Research with ALS data using deep learning methods has involved spatially simple objects, e.g., burial mounds [1–3,20,21], charcoal piles [1,3,21,22], post-mining pits [23] as well as roundhouses and huts [2]; attempts to identify more spatially complex structures such as polygonal megalithic tombs [24] were also undertaken. Deep learning methods have also been implemented for the detection of multi-element objects with complex, irregular relief such as Celtic field systems [1,20].

The following deep neural network models were applied in the recognition of the archaeological monuments based on the ALS data: R-CNN (Region Based Convolutional Neural Networks) [21], Faster-R-CNN [1,20], Mask R-CNN [24] and U-Net [24]. The following feature extraction modules were used: ResNet18 [2], ResNet50 [20], ResNet101 [21] and VGG16 (Visual Geometry Group) [1,25].

Conventional ALS data acquisition and processing is usually performed according to the following procedures: (1) acquisition of data, (2) classification of the point cloud, (3) generation of a DTM (Digital Terrain Model) followed by transformation into a visualisation (4). The preferred method of ALS data transformation to increase the visibility of the detected objects is LRM (Local Relief Model) [1–3,20,22,23]. This technique is used to extract small changes in terrain relief and is based on subtracting the mean values from the input DTM [26]. Machine learning has also been used to analyse archaeological resources directly from a point cloud [27].

1.3. Celtic Fields and Burial Mounds

Prehistoric and ancient field systems are among the best-preserved relics of the past agricultural landscape in western, central and northern Europe. The commonly used name "Celtic fields" was introduced by O.G.S. Crawford, who first described them in southern England in 1923 [28]. Celtic fields are composed of linear terrain elevations formed of soil and stones (from now on referred to as *field banks*) enclosing quadrangular or square land areas that might have been cultivated (further on referred to as *field plots*). Those past agricultural systems are characterized by a recognisable spatial pattern and can be traced chronologically from the Neolithic to the Iron Age [29–40]. The field banks can be a few centimetres to about a meter high and a few meters to over a dozen meters wide. The length of the banks varies ranging from a few meters to over a hundred meters. Within the enclosures, denivelations with a regular linear pattern may be preserved, which are most likely the remains of former agricultural earthworks. A field system may cover large areas of up to 200 hectares [35]. Several studies have indicated that Panicum miliaceum, Hordeum vulgare, Triticum dicoccum, Triticum spelta and Avena sativa were cultivated there [34,41–43]. Ancient field systems can be easily irreparably damaged, for example by scarification or deep ploughing. Nevertheless, many of them can still be detected in modern forests. Due to their location in areas with favourable agricultural conditions, some of the fields had been reused and formed a mosaic of cultivated areas in successive periods of use.

In Poland there are relics of ancient agricultural field systems with different spatial patterns. Field systems whose spatial arrangement is consistent with Celtic fields described in other European countries have also been identified. Using ALS data from forested areas, Banasiak and Berezowski discovered hundreds of such locations over the whole area of country via desk-based work. The discovery was supported by AlexNet and ResNet [44]. Despite extensive occurrence of Celtic fields, excavations and dating have been very limited. Approximately 300 km of linear terrain elevations of earth and mudstone of 2–8 m in width and several dozens of centimetres in height were identified in many areas of the investigated complex of the Białowieża Forest. Fragments of the discovered field systems were investigated by excavation as well as geophysical and soil analyses. This made it possible to date them back to the pre-Roman period, the period of Roman influence and the early Middle Ages [7,45–48]. Remains of ancient field systems chronologically associated with the Late Roman period have also been detected in the Tuchola Forest [49].

In the Białowieża Forest there are various mounds of anthropogenic origin. Until the remains of ancient field systems were found, the most numerous group of archaeological monuments in this area had been burial mounds. The latter are individual or group forms; some of these were investigated by excavations and dated back to the period between the 1st and the 12th century AD [50–58]. They are typically round shape but vary in size. As previous research has shown, a large group of mounds are structures related to historical processing of raw wood, including charcoal piles, or sites for the production of potash and tar [55,59].

2. Materials and Methods

2.1. Characteristics of the Area—The Primeval Forest of Białowieża

The Białowieża Forest (BF) is one of the last and largest remnants of a vast primeval forest that once stretched across the European lowlands. It is characterised by a high degree of biodiversity, including species-rich, multi-layered forest stands of varying ages. The topography is relatively flat (131.6–195.6 m above sea level). The Białowieża National Park (BNP) and nature reserves are protected by both national and international law (UNESCO) (Figure 1).



Figure 1. Research area—the Polish part of the Białowieża Forest with the training area marked in red.

The character of primeval forest, mainly located in the BNP, and the modern legal protection of the best-preserved fragments of the BF make it an area with limited modern and historical human activities. This could be the reason for such a large number of anthropogenic relics, which, under other circumstances, could have been destroyed or altered.

2.2. Airborne Laser Scanning and Digital Terrain Model

An airborne laser scanning mission was conducted during the leaf-off season between 25 November and 7 December 2015. Acquiring data during the leaf-off season allowed better mapping of the area under the tree canopy. The ALS dataset was collected using a Riegl LMS-Q680i full waveform system (installed on the Vulcanair P-68 Observer platform) at an average speed of 105 km/h and from an altitude of 500 m above ground level (AGL). The average point density was 11 points/m². The horizontal and vertical accuracies of the acquired point cloud were 0.20 m and 0.15 m, respectively.

The Digital Terrain Model was calculated using ALS point cloud. Only LiDAR points classified as ground—2nd class ASPRS LAS (American Society for Photogrammetry and Remote Sensing LASer) were used. Finally, a grid with a resolution of 0.50 m was created using the Triangular Irregular Network (TIN) method.

2.3. Digital Terrain Model Visualizations

The Relief Visualization Toolbox ver. 2.2.1 (https://iaps.zrc-sazu.si/ accessed on 6 February 2022) [60] was used to create the DTM visualization. SLRM (Simple Local Relief Model—radius for trend assessment 20 pixels, 8 bits), LD (Local Dominance—minimum/maximum pixel radius 10/20) and an analytical hillshading (Z-factor = 4) were created.

Images of the same area obtained with three different visualization techniques were combined into one image referred to as modality fusion. The channels of the RGB colour model (red, green, blue) were filled with SLRM, analytical hillshading and LD images, respectively (Figure 2). A combination of different DTM visualizations allows the mutual complement of information between images. In the next step, the obtained raster data

were transferred to the QGIS ver. Bucaresti (QGIS Development Team. QGIS Geographic Information System, https://qgis.org accessed on 6 February 2022).

Figure 2. Examples of DTM visualization (250×250 m): SLRM, analytical hillshading, LD, modality fusion.

2.4. Research and Training Area

The research area is located in the territory of the Browsk, Hajnówka and Białowieża Forest Districts and the Białowieża National Park, i.e., the Polish part of the Białowieża Primeval Forest, and its total area is 697.8 km² (Figure 1). The research dataset includes the DTM area obtained within the Life+ ForBioSensing project.

The choice of location for the training area was based on the best possible representation of the objects from the listed classes (described in Section 2.6). The training area has a rectangular shape, a size of 5×10 km (50 km²) and is located on the territory of the Białowieża Forest District and the Białowieża National Park (Figure 1).

2.5. Research and Training Datasets

The research dataset is a collection of 11,165 tiles containing the fused DTM visualizations described above. Each image represents an area 250×250 m ALS data coverage and is georeferenced. The research dataset was created to obtain a semantic segmentation mask for the whole area.

The training dataset is used for learning the model and consists of pairs of input images and the corresponding ground truth masks (Figure 3). The software included QGIS, GIMP (GNU Image Manipulation Program, https://www.gimp.org/ accessed on 6 February 2022) and Gimp Selection Feature Plugin QGIS (Luiz Motta, https://plugins.qgis.org/plugins/gimpselectionfeature_plugin/ accessed on 6 February 2022).

The training datasets are split into the training set, validation set and test set. The training area was divided into 200 square areas of 500 by 500 m, of which a training set (160 tiles) and a test set (40 tiles) were randomly derived. The validation set was derived from the training set (see Section 2.7.2). For each tile obtained in this way, 13 images sized 256×256 m (512 pixels) were generated, as shown in Figure 4. The database has separate areas for the training set and the test set. The tiles of the training set and the validation set overlap to produce unique images from one area. By using the overlapping regions technique, the database could be increased by more than three times.



Figure 3. Manually drawn ground truth mask of the training area.



Figure 4. Training dataset preparation—square land section of the training area (5×10 kilometres) with randomly generated train/validation sets (green) and test set (blue). Points are the centroids of the generated tiles. Enlarged fragment (red box) and single training image (yellow box with dot).

2.6. Classes of the Objects

In the initial phase, deep learning neural network experiments focused on two classes of archaeological objects, namely, field systems banks and burial mounds. Low intersection over union (IoU) values and a large number of false positives were observed, mainly due to misidentified inland dunes. Therefore, the set of classes was expanded by inland dunes to include objects with competing shapes and textures. Based on a characteristic terrain relief, a novel class was added to archaeological features, defined as field system plots.

The burial mounds class includes 211 objects in the training area. The class was created based on visual resemblance to the burial mounds excavated and described in recent decades. Examples of objects from this class are shown in Figure 5. Finally, 9 classes of objects, including archaeological, natural and modern objects, listed in Table 1, were prepared for training the U-Net model.



Figure 5. Examples $(30 \times 30 \text{ m})$ from the burial mounds category. The visualization used is the Simple Local Relief Model.

Archaeological features:		
1	field system banks	
2	field system plots	
3	burial mounds	
Modern features:		
4	roads	
5	forest paths and divisions	
6	modern landscape (e.g., houses, farmlands)	
Natural features:		
7	inland waterways	
8	inland dunes	
Remaining land/area:		
9	background	

Table 1. List of classes used to train the U-Net model.

2.7. Deep Learning—Method Description and Training

2.7.1. Data Description

The prepared dataset consisted of 2600 images comprised 512×512 pixels, the basis for deriving the training and test sets. The training set and test set consisted of 2080 and 520 images, respectively; each has an associated mask with regions of interest; each class was marked with a specific colour. The data were distinguished between 9 classes (including



background). Figure 6 shows an example pair of an input image and a corresponding mask (Celtic fields).

Figure 6. Example pair of an image and its ground truth mask (Celtic fields).

2.7.2. Data Preparation

In the data preparation phase, both the images and the masks were divided into 256×256 -pixel patches. Due to class imbalance (Figure 7) and one class domination in the training set, the patches that contained only pixels assigned to the background class were removed. Then the fields were split into the training and validation datasets with a 0.2 (rate) validation split. The samples in the training dataset were also enlarged. The ground truth image and mask were expanded to obtain 5 additional samples after applying various combinations of the following operations: horizontal flipping, vertical flipping and grid distortion.



Figure 7. Histogram showing the pixel-wise distribution of training set.

Augmentations were performed using an augmentation library [61]. The prepared sets were normalized to be in the range of 0–1. As a final step in the preprocessing phase, the RGB masks were encoded using a one-hot encoding, also known as the "dummy variable" method. This method also ensures normalization of categorical variables. The encoding transforms the label tensor from the form of (N,W,H,3) to (N,W,H,9), where N is the number of samples, W is the width and H is the height; the last parameter represents a change



from the number of channels to the number of classes. The final dataset consisted of 7779 training and 1945 validation patches (Figure 8).

Figure 8. Histogram showing the distribution of training, validation and test set.

2.7.3. Problem Description

Semantic segmentation is the pixel-by-pixel labelling of an image and has recently emerged as one of the fundamental tasks that learning algorithms attempt to solve. With the advent of artificial neural networks, especially convolutional neural networks, the challenge has recently been solved using state-of-the-art (SOTA) techniques. Ulku and Akagunduz [62] listed these methods and summarized their performance and limitations.

2.7.4. Model Architecture

For this project, we have chosen an encoder-decoder approach based on the U-NET architecture (Figure 9) [19]. The encoder part is responsible for feature extraction and data dimensionality reduction with sequentially stacked convolutional layers and max-pooling layers. The decoder part is built with the upsampling and convolutional layers to create a probability mask that describes objects in the image. The feature maps generated during the encoding phase are pooled with the feature maps generated by the decoder to improve and speed up the learning process.

Different types of convolutional neural networks were tested for the encoder. The best results were obtained with ResNet-34, a 34-layer convolutional neural network [16]. The weights for this network were not initialized randomly. Instead, we used a pre-trained model on the ImageNet dataset as initial weights. In addition, the batch normalization layers were added to the decoder part to further stabilise the training process. The final model had a total of 25,679,458 parameters. The model was developed using the Keras Deep Learning library (Chollet, F., and others. (2015). Keras. GitHub. Retrieved from https://github.com/fchollet/keras accessed on 6 February 2022).



Figure 9. Example U-Net architecture. Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on the top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. Arrows denote functions [19].

2.7.5. Training

The developed model was trained on an Nvidia RTX 2080Ti (NVIDIA, USA) graphics card, with 11GB of video-RAM memory (Random-Access Memory). The training was carried out using 'Adam', a first-order, gradient-based optimiser [63]. A set of callbacks was defined to effectively train the model and save its results:

Model checkpoint—allowed saving model weights that achieved the best results in the validation set.

Reduce learning rate on plateau—implemented a decaying learning rate dropped by a decay factor of 0.5 if the metrics score did not improve for 5 succeeding epochs.

Two metrics were applied to measure model performance effectively, namely:

Intersection over Union, also referred to as the Jaccard index—measured the overlap between generated and ground truth masks for each of the defined classes. The mean score across all classes was also calculated. It was defined as follows:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

Dice-Sorensen coefficient/ F1-score—similar to the IoU score; there is a positive correlation between the two. Its metric definition is as follows:

$$DS(A,B) = \frac{2*|A \cap B|}{|A|+|B|}$$
(2)

For both metrics, the highest possible value is 1.0 and the lowest value is 0.0.

To measure the error and adjusting weights in the most effective way, a combination of two loss functions was used:

Jaccard loss—loss function based on the IoU/Jaccard index; first introduced by Bertels et al. [64].

Categorical focal loss—due to class imbalance in the dataset, focal loss was introduced to handle this problem [65].

The two losses were added to determine the final loss subsequently used to train the model.

3. Results

The model was trained on 100 epochs. However, since both metrics and loss functions reached a plateau, the process was stopped at the 83rd epoch. The best results were obtained with the following hyperparameters: Learning Rate—0.001 (start), Batch Size—32 samples, and Epochs—100. The learning curves for the training set (orange), validation set (blue), and learning rate decay levels are shown below (Figures 10–13).



Figure 10. F1-score learning curve.



Figure 11. IoU learning curve.



Figure 12. Loss curve.



Figure 13. Learning rate decay.

Table 2 shows metrics and loss values that the model reached at its best.

Table 2. F1-score, IoU and Loss scores for the training, validation and test set.

	F1-Score	IoU	Loss
Training set	0.91	0.84	0.16
Validation set	0.79	0.68	0.33
Test set	0.58	0.50	0.53

The values of the IoU metric for the included classes on the test set are presented in Table 3. The IoU metric was 0.41 for ancient field system banks, 0.616 for ancient field system plots and 0.62 for burial mounds.

Table 3. IoU metric results for the test set.

Class	IoU
field system banks	0.408
field system plots	0.616
burial mounds	0.615
roads	0.673
forest paths and divisions	0.514
modern landscape	0.531
inland waterways	0.573
inland dunes	0.333
background	0.782

Examples of results in the analysed object classes are shown in Figure 14 (from the left: input image, ground truth mask and prediction masks). The result of the network operations seems to fulfil the assumptions of the study. The obtained segmentation masks are characterized by good recognition of the objects' contours (Figure 14A–J).



Figure 14. Test set tiles samples (**A**–**J**), manually drawn ground truth mask and the results of semantic segmentation (prediction masks) of the test area with the U-Net model.

It should be noted that it is almost impossible to obtain a mask of an object's true shape in the class of field system's banks since it is based on subjective human assessment. The model repeatedly made more accurate decisions, e.g., detecting gaps in the banks of the field system (Figure 14B,C). The above-mentioned objects' IoU has resulted from the recognition system errors and human errors.

In the next phase, the trained U-Net model was used to perform semantic segmentation of the research dataset to obtain predicted segmentation masks. The latter were then georeferenced and spatially aggregated to produce a single terrain segmentation map. The smaller scale representation of the results, shown in Figure 15, is the area of the Białowieża National Park covering over 105 km² (10,558 ha).

In the Białowieża Forest, the ancient field systems occur in the form of clusters of various sizes. The model recognized field systems at all previously known sites, and, in most cases, the quality of the preserved mask was satisfactory (Figures 14, 15 and 16A–D). However, some sites exhibited fragmentary classification of the banks (Figure 16E). In these cases, low quality of the mask was much more related to lack of detection (false negatives) than to false positives, the most probable reason being the insufficiently representative training set.

Inland dunes proved to be the main cause of false positives in our previous experiments; these objects were classified by the neural networks as field system banks. Increasing the number of training classes to include the objects with competing visual features and applying a fusion of three DTM transformations lessened this problem. Only few false positives occurred outside the fields' area (Figure 16).

The class of ancient field system plots was defined based on the characteristic terrain relief within the land enclosures. The IoU metric of 0.616 seems to confirm this identification.

In the burial mound class, most objects were correctly marked (IoU = 0.615), but there were also areas with incomplete recognition. Small diameter burial mounds were sometimes ignored by the neural network (Figure 17A,B). In addition, numerous misdiagnoses (noise) with a very small area (approx. 0.5 m^2) were made (Figure 17C,D). The same problem, although less pronounced, concerned the other classes.



Figure 15. Semantic segmentation of the ALS data from the area of the Białowieża National Park by U-Net model.



Figure 16. Example of ancient field system prediction mask contour (in red) from 5 areas of the Białowieża Forest (**A**–**E**). (**F**) is a map of the research area with the location of the examples.



Figure 17. Examples of problems in the segmentation of objects in the burial mounds class (blue outline): (**A**,**B**)—incomplete recognition; (**C**,**D**)—misdiagnoses/noises visible on a large scale.

4. Discussion

The application of the deep neural model U-Net to ALS data obtained from the Polish part of the Białowieża Forest helped identify two types of archeological monuments: (1) simple structures (round mounds)—burial mounds; and (2) complex structures with multiple elements—ancient field systems. The latter were divided into two classes: (A) linear land elevations called banks, which form enclosures; and (B) potentially cultivated areas called plots.

The U-Net deep learning network proved very useful in semantic segmentation of multimodal images derived from ALS data. To evaluate the performance of the model, two metrics were used in our research, i.e., Dice coefficient (DSC)/F1 score (F1) and Intersection over Union (IoU). The F1 metric for the analysed object classes reached 0.91 for the training dataset and 0.58 for the test dataset. The Intersection-over-Union (IoU) metric [23], also known as the Jaccard index, was used to measure the overlap between the predicted segmentation mask and ground truth mask. The segmentation accuracy of the U-Net model in the classes of ancient field system banks, ancient field system plots and burial mounds achieved IoU values of 0.41, 0.62 and 0.62 in the test dataset. These results cannot be directly compared with those of other authors since previous research used different tasks and objectives. The WODAN2 (Workflow for Object Detection of Archeology in the Netherlands) system, based on the Faster R-CNN model, has been described as a useful analytical tool for object detection tasks [1]. For the categories of Celtic fields and barrows, the model achieved an efficiency of 46% and 50%, respectively. Knowledge of the classification problem in the WODAN2 study: false-positive object classification in the case of inland dunes, and our previous experience, caused us to expand the list of classes to include objects that had been misclassified in previous attempts. The U-Net architecture has already been used in the semantic segmentation of archaeological objects and shows an advantage over Mask R-CNN in the identification of such objects of, e.g., ancient Maya civilization [24].

The class of ancient field system plots was included in the dataset after significant differences in the terrain relief of these areas had been identified. The purpose of this step was to test the ability of deep learning methods to detect (feature extraction) gentle terrain relief details, which are not limited to linear terrain elevations (banks). The U-Net output segmentation mask overlaying the ground truth mask reached an IoU value of 0.616 in this class. The result still needs to be confirmed using larger datasets in other locations. However, this solution, i.e., inclusion of ancient field system plots class, might enhance ancient fields detection.

Most researchers applying deep learning techniques to ALS archaeological data use DTM visualized as LRM. Such conversion provides a single-channel grayscale image while most deep learning models expect a 3-channel RGB image as input. Hence, either the neural network input must be changed [20] or the data must be tripled to fill all RBG channels [2]. To improve the visualization of different objects, the DTM data can be transformed in many ways. In our study, three DTM transformations (SLRM, shaded image and LD) that had been found optimal for archaeological object detection were fused into one image based on the RGB model. This solution increased the efficiency of the system for the recognition of historical objects already in the phase of data preparation. Other authors also reviewed the usefulness of different DTM visualizations in the field of computer vision and found no clear, unambiguous advantage of one visualization over another [66].

The main source of problems in creating the database were the objects that were difficult to distinguish. The banks of the ancient field system have poorly visible boundaries/contours that are seamless and continuous with the surrounding environment. Object labelling is the most time-consuming step in creating the database. The use of various DTM visualizations and software graphics programs helps improve this process. Despite the inaccuracy of the training mask resulting from human errors, our trained model detected correct outlines of objects; breaks in banks were also more accurately indicated. Similar observations were made by other researchers [21]. The ResNet-FPN-101 model was better at recognizing the true shape of an object than the ground truth mask. Archaeological inventories conducted with the use of deep learning methods can significantly reduce the workload and error rate of traditional archaeological desk-based surveys [67].

The proposed method provides an opportunity to improve analytical work and more efficiently verify the results of a standard archaeological inventory, mainly by reducing the time for large-scale acquisition. Research has shown that such solutions can help assess the density of archaeological objects over large areas and complex terrain [68]. The limitation of machine learning methods seems to include different quality of ALS data (density of land cover by survey, blind spots, etc.) and the condition of historical objects [67]. Another weakness of machine learning is the tendency to ignore rare and unusual objects as the models cannot fully account for the variability of the classified features and operate on a limited set of pre-selected variables. Hence, they lack information necessary to differentiate archeologically relevant classes [69]. The number of training images for U-Net deep neural network is limited in this work but it can be easily extended to achieve better generalization abilities. It should also be noted that, considering the effort and time needed to create ground truth masks, a tool should be sought for automatic creation thereof.

5. Conclusions

The implemented machine learning solution allows automatic segmentation of the transformed ALS data. The tool supports the process of detection and mapping of the cultural resources of extensive areas. The developed system is a complete functional tool ready for the number of classes to be increased and retraining with new datasets.

Author Contributions: P.Z.B., P.L.B. and K.D. played a critical part in the preparation of the manuscript. Conceptualization, P.Z.B. and P.L.B.; Methodology, P.Z.B., P.L.B. and K.D.; Software, P.Z.B. and K.D.; Validation, P.Z.B., P.L.B. and K.D.; Formal analysis, P.Z.B. and K.D.; Investigation, P.Z.B., P.L.B. and K.D.; Resources, K.S., P.Z.B., P.L.B. and M.M.; Data curation, P.Z.B., P.L.B. and K.D.; Writing—original draft, P.Z.B. and P.L.B.; Writing—review and editing, K.D., K.S., M.M. and R.Z.; Visualization, P.Z.B., K.D. and M.M.; Supervision, P.Z.B. and K.S.; Project administration, P.Z.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Airborne Laser Scanner data were collected as part of the project LIFE + For-BioSensing PL Comprehensive monitoring of stand dynamics in Białowieża Forest supported with remote sensing techniques co-funded by the European Commission Life+ program (contract number LIFE13 ENV/PL/000048) and the National Fund for Environmental Protection and Water Management in Poland (contract number 485/2014/WN10/ OP-NM-LF/D).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Verschoof-van der Vaart, W.B.; Lambers, K.; Kowalczyk, W.; Bourgeois, Q.P.J. Combining Deep Learning and Location-Based Ranking for Large-Scale Archaeological Prospection of LiDAR Data from The Netherlands. *ISPRS Int. J. Geo Inf.* 2020, *9*, 293. [CrossRef]
- Trier, Ø.D.; Cowley, D.C.; Waldeland, A.U. Using Deep Neural Networks on Airborne Laser Scanning Data: Results from a Case Study of Semi-automatic Mapping of Archaeological Topography on Arran, Scotland. *Archaeol. Prospect.* 2018, 26, 165–175. [CrossRef]
- 3. Lambers, K.; Verschoof-van der Vaart, W.; Bourgeois, Q. Integrating Remote Sensing, Machine Learning, and Citizen Science in Dutch Archaeological Prospection. *Remote Sens.* **2019**, *11*, 794. [CrossRef]
- 4. Tadeusiewicz, R.; Szaleniec, M. Leksykon Sieci Neuronowych; Wydawnictwo Fundacji "Projekt Nauka": Wrocław, Poland, 2015; p. 94.
- 5. Kurczyński, Z. Airborne Laser Scanning in Poland—Between Science and Practice. Archives of Photogrammetry. *Cartogr. Remote Sens.* **2019**, *31*, 105–133. [CrossRef]
- 6. Kurczyński, Z.; Bakuła, K. Generowanie referencyjnego numerycznego modelu terenu o zasięgu krajowym w oparciu o lotnicze skanowanie laserowe w projekcie ISOK. In *Monografia "Geodezyjne Technologie Pomiarowe": Wydanie Specjalne*; Kurczyński, Z., Ed.; Zarząd Główny Stowarzyszenia Geodetów Polskich: Warsaw, Poland, 2013; pp. 59–68.
- Stereńczak, K.; Zapłata, R.; Wójcik, J.; Kraszewski, B.; Mielcarek, M.; Mitelsztedt, K.; Białczak, M.; Krok, G.; Kuberski, Ł.; Markiewicz, A.; et al. ALS-Based Detection of Past Human Activities in the Białowieża Forest—New Evidence of Unknown Remains of Past Agricultural Systems. *Remote Sens.* 2020, 12, 2657. [CrossRef]
- 8. Zapłata, R.; Stereńczak, K.; Kraszewski, B. Wielkoobszarowe badania dziedzictwa archeologicznego na terenach leśnych. *Kur. Konserw.* **2018**, *15*, 47–51.
- 9. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Comput. Intell. Neurosci.* **2018**, 2018, 7068349. [CrossRef]
- 10. LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput.* **1989**, *1*, 541–551. [CrossRef]
- 11. LeCun, Y.; Haffner, P.; Bottou, L.; Bengio, Y. Object recognition with gradient-based learning. In *Shape, Contour and Grouping in Computer Vision*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 319–345. [CrossRef]
- 12. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
- Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* 2017, 60, 84–90. [CrossRef]
- 14. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the 3rd International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
- Szegedy, C.; Wei, L.; Yangqing, J.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; 2015. [CrossRef]
- 16. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; 2016. [CrossRef]
- 17. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015. [CrossRef]
- Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, 2481–2495. [CrossRef] [PubMed]

- Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lect. Notes Comput. Sci.* 2015, 234–241. [CrossRef]
- 20. Verschoof-van der Vaart, W.B.; Lambers, K. Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands. *J. Comput. Appl. Archaeol.* **2019**, *2*, 31–40. [CrossRef]
- Kazimi, B.; Thiemann, F.; Sester, M. Object Instance Segmentation in Digital Terrain Models. In Proceedings of the Computer Analysis of Images and Patterns, Salerno, Italy, 3–5 September 2019; pp. 488–495. [CrossRef]
- 22. Bakuła, K.; Ostrowski, W.; Zapłata, R. Automatyzacja w Procesie Detekcji Obiektów Archeologicznych z Danych ALS. Folia Praehist. Posnaniensia 2014, 19, 189. [CrossRef]
- Gallwey, E.; Tonkins, C. Bringing Lunar LiDAR Back Down to Earth: Mapping Our Industrial Heritage through Deep Transfer Learning. *Remote Sens.* 2019, 11, 1994. [CrossRef]
- 24. Bundzel, M.; Jaščur, M.; Kováč, M.; Lieskovský, T.; Sinčák, P.; Tkáčik, T. Semantic Segmentation of Airborne LiDAR Data in Maya Archaeology. *Remote Sens.* 2020, 12, 3685. [CrossRef]
- Politz, F.; Kazimi, B.; Sester, M. Classification of laser scanning data using deep learning. In Proceedings of the 38th Scientific-Technical Annual Conference of the DGPF and PFGK18, Munich, Germany, 6–9 March 2018.
- Hesse, R. LiDAR-Derived Local Relief Models—A New Tool for Archaeological Prospection. Archaeol. Prospect. 2010, 17, 67–72. [CrossRef]
- Arnold, N.; Angelov, P.; Viney, T.; Atkinson, P. Automatic Extraction and Labelling of Memorial Objects From 3D Point Clouds. J. Comput. Appl. Archaeol. 2021, 4, 79–93. [CrossRef]
- 28. Crawford, O.G.S. Air Survey and Archaeology. Geogr. J. 1923, 61, 342. [CrossRef]
- Arnoldussen, S.; van der Linden, M. Palaeo-Ecological and Archaeological Analysis of Two Dutch Celtic Fields (Zeijen-Noordse Veld and Wekerom-Lunteren): Solving the Puzzle of Local Celtic Field Bank Formation. *Veg. Hist. Archaeobotany* 2017, 26, 551–570. [CrossRef]
- Arnold, V. Älter als die Römer: Bisher übersehene Spuren einstiger 30, Beackerung unter bayerischen Wäldern. Forstl. Forsch. München 2020, 218, 8–18.
- Klamm, M. Aufbau und Entstehung eisenzeitlicher Ackerfluren ("celtic fields") Neue Untersuchungen im Gehege Ausselbek, Kr. Schleswig-Flensburg. Archäologische Unters. 1993, 16, 122–124.
- 32. Spek, T.; Waateringe, W.G.; Kooistra, M.; Bakker, L. Formation and Land-Use History of Celtic Fields in North-West Europe—An Interdisciplinary Case Study at Zeijen, the Netherlands. *Eur. J. Archaeol.* **2003**, *6*, 141–173. [CrossRef]
- Zimmerman, W.H. Die eisenzeitlichen Ackerfluren–Typ 'Celtic field'–von Flögeln-Haselhörn, Kr. Wesermünde. Probl. Der Küstenforschung Im Südlichen Nordseegebiet 1976, 11, 79–90.
- 34. Kooistra, M.J.; Maas, G.J. The Widespread Occurrence of Celtic Field Systems in the Central Part of the Netherlands. *J. Archaeol. Sci.* 2008, *35*, 2318–2328. [CrossRef]
- 35. Meylemans, E.; Creemers, G.; De Bie, M.; Paesen, J. Revealing extensive protohistoric field systems through high resolution LIDAR data in the northern part of Belgium. *Archäologisches Korrespondenzblatt* **2015**, *45*, 1–17.
- 36. Nielsen, V. Prehistoric field boundaries in Eastern Denmark. J. Dan. Archaeol. 1984, 3, 135–163. [CrossRef]
- 37. Arnold, V. »Celtic Fields« and Other Prehistoric Field Systems in Historical Forests of Schleswig-Holstein from Laser-Scan Dates. *Archäologisches Korrespondenzblatt* 2011, 41, 439–455. [CrossRef]
- Løvschal, M. Lines of Landscape Organization: Skovbjerg Moraine (Denmark) in the First Millennium BC. Oxf. J. Archaeol. 2015, 34, 259–278. [CrossRef]
- 39. Whitefield, A. Neolithic 'Celtic' Fields? A Reinterpretation of the Chronological Evidence from Céide Fields in North-Western Ireland. *Eur. J. Archaeol.* 2017, 20, 257–279. [CrossRef]
- 40. Behre, K.E. Zur Geschichte der Kulturlandschaft Nordwestdeutschlands seit dem Neolithikum. *Ber. Der Römisch Ger. Komm.* 2002, 83, 39–68.
- 41. Buurman, J. Graan in Ijzertijd-silos uit Colmschate in Voordrachten gehouden te Middelburg ter gelegenheid van het afscheid van Ir. JA Trimpe Burger als provinciaal archeoloog van Zeeland. *Ned. Archeol. Rapp.* **1986**, *3*, 67–73.
- Vermeeren, C. Cultuurgewassen en onkruiden in Ittersumerbroek. In Bronstijdboeren in Ittersumerbroek. Opgraving van een Bronstijdnederzetting in Zwolle-Ittersumerbroek; Clevis, H., Verlinde, A.D., Eds.; Stichting Archeologie IJssel/Vechtstreek: Kampen, The Netherlands, 1991; pp. 93–106.
- 43. Bakels, C.C. Fruits and seeds from the Iron Age settlements at Oss-Flussen. Analecta Praehist. Leiden. 1998, 30, 338–348.
- 44. Banasiak, P.; Berezowski, P. Ancient Fields Systems in Poland: Discovered by Manual and Deep Learning Methods. Available online: https://prohistoric.wordpress.com (accessed on 31 December 2018).
- 45. Zapłata, R.; Stereńczak, K. Archeologiczna niespodzianka w Puszczy Białowieskiej. Archeol. Żywa 2017, 1, 63.
- Zapłata, R.; Stereńczak, K. Puszcza Białowieska, LiDAR i dziedzictwo kulturowe–zagadnienia wprowadzające. *Raport* 2016, 11, 239–255.
- 47. Stereńczak, K.; Krasnodębski, D.; Zapłata, R.; Kraszewski, B.; Mielcarek, M. Sprawozdanie z Realizacji Zadania Inwentaryzacja Dziedzictwa Kulturowego, w Projekcie pt. Ocena Stanu Różnorodności Biologicznej w Puszczy Białowieskiej na Podstawie Wybranych Elementów Przyrodniczych i Kulturowych; Forest Research Institute: Sękocin Stary, Poland, 2016.

- 48. Zapłata, R.; Stereńczak, K.; Grześkowiak, M.; Wilk, A.; Obidziński, A.; Zawadzki, M.; Stępnik, J.; Kwiatkowska, E.; Kuciewicz, E. Raport Końcowy. Zadanie "Inwentaryzacja Dziedzictwa Kulturowego" na Terenie Polskiej części Puszczy Białowieskiej w Ramach Projektu "Ocena i Monitoring Zmian Stanu Różnorodności Biologicznej W Puszczy Białowieskiej Na Podstawie Wybranych Elementów Przyrodniczych I Kulturowych–Kontynuacja"; Forest Research Institute: Sękocin Stary, Poland, 2019.
- 49. Sosnowski, M.; Noryśkiewicz, A.M.; Czerniec, J. Examining a scallop shell-shaped plate from the Late Roman Period discovered in Osie (site no.: Osie 28, AZP 27-41/26), northern Poland. *Analecta Archaeol. Ressoviensia* **2019**, *14*, 91–98. [CrossRef]
- 50. Górska, I. Badania archeologiczne w Puszczy Białowieskiej. Archeol. Pol. 1976, 21, 109–134.
- Götze, A. Archäologische untersuchungen im urwalde von bialowies. In Beiträge zur Natur- und Kulturgeschichte Lithauens und Angrenzender Gebiete; Stechow, E., Ed.; Verlag der Bayerischen Akademie der Wissenschaften in Kommission des Verlags R. Oldenburg München: Munich, Germany, 1929; pp. 511–550.
- 52. Oszmiański, M. Inwentaryzacja Kurhanów na Terenie Puszczy Białowieskiej; Wojewódzki Urząd Ochrony Zabytków w Białymstoku: Białystok, Poland, 1996.
- 53. Krasnodębski, D.; Olczak, H. Badania archeologiczne na terenie polskiej części Puszczy Białowieskiej—Stan obecny, problemy i perspektywy. In *Biuletyn Konserwatorski Województwa Podlaskiego*; Wojewódzki Urząd Ochrony Zabytków w Białymstoku: Białystok, Poland, 2012; Volume 18, pp. 145–168.
- 54. Krasnodębski, D.; Olczak, H. Puszcza Białowieska jako przykład badań archeologicznych na obszarach leśnych—Wyniki i problemy przeprowadzonej w 2016 r. inwentaryzacji dziedzictwa kulturowego. *Podl. Zesz. Archeol.* **2017**, *13*, 5–64.
- 55. Zapłata, R.; Wilk, A.; Grześkowiak, M.; Obidziński, A.; Zawadzki, M.; Stereńczak, K.; Kuberski, Ł. Sprawozdanie Końcowe w Związku z Realizacją Umowy nr EO.271.3.5.2019 z Dnia 29 Marca 2019 r. "Dziedzictwo Kulturowe i Rys Historyczny Polskiej części Puszczy Białowieskiej"; Dyrekcja Generalna Lasów Państwowych: Warsaw, Poland, 2019.
- 56. Samojlik, T. Antropogenne Przemiany Środowiska Puszczy Białowieskiej do Końca XVIII Wieku; Zakład Badania Ssaków PAN: Białowieża-Kraków, Poland, 2007.
- 57. Krasnodębski, D.; Olczak, H.; Mizerka, J.; Niedziółka, K. Alleged burial mounds from the late Roman Period at Leśnictwo Sacharewo site 3, Białowieża Primeval Forest. Światowit **2018**, *57*, 89–99. [CrossRef]
- 58. Olczak, H.; Krasnodębski, D.; Szlązak, R.; Wawrzeniuk, J. The Early Medieval Barrows with Kerbstones at the Leśnictwo Postołowo Site 11 in the Białowieża Forest (Szczekotowo Range). *Spraw. Archeol.* **2020**, 72. [CrossRef]
- 59. Samojlik, T. Rozkwit i upadek produkcji potażu w Puszczy Białowieskiej w XVII-XIX w. Rocz. Pol. Tow. Dendrol. 2016, 64, 9–19.
- 60. Kokalj, Ž.; Hesse, R. Airborne Laser Scanning Raster Data Visualization; Založba ZRC: Ljubljana, Slovenia, 2017. [CrossRef]
- 61. Buslaev, A.; Iglovikov, V.I.; Khvedchenya, E.; Parinov, A.; Druzhinin, M.; Kalinin, A.A. Albumentations: Fast and Flexible Image Augmentations. *Information* **2020**, *11*, 125. [CrossRef]
- 62. Ulku, I.; Akagunduz, E. A survey on deep learning-based architectures for semantic segmentation on 2D images. *arXiv* 2019, arXiv:1912.10230. [CrossRef]
- 63. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. arXiv 2014, arXiv:1412.6980.
- 64. Bertels, J.; Eelbode, T.; Berman, M.; Vandermeulen, D.; Maes, F.; Bisschops, R.; Blaschko, M.B. Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice. *Lect. Notes Comput. Sci.* **2019**, 92–100. [CrossRef]
- Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 2020, 42, 318–327. [CrossRef]
- Somrak, M.; Džeroski, S.; Kokalj, Ž. Learning to Classify Structures in ALS-Derived Visualizations of Ancient Maya Settlements with CNN. *Remote Sens.* 2020, 12, 2215. [CrossRef]
- Bonhage, A.; Eltaher, M.; Raab, T.; Breuß, M.; Raab, A.; Schneider, A. A Modified Mask Region-based Convolutional Neural Network Approach for the Automated Detection of Archaeological Sites on High-resolution Light Detection and Ranging-derived Digital Elevation Models in the North German Lowland. *Archaeol. Prospect.* 2021, 28, 177–186. [CrossRef]
- 68. Davis, D.S. Defining What We Study: The Contribution of Machine Automation in Archaeological Research. *Digit. Appl. Archaeol. Cult. Herit.* **2020**, *18*, e00152. [CrossRef]
- 69. Bickler, S.H. Machine Learning Arrives in Archaeology. Adv. Archaeol. Pract. 2021, 9, 186–191. [CrossRef]