Journal of Cultural Heritage xxx (xxxx) xxx–xxx

Available online at
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

Original article

Using CarcassonNet to automatically detect and trace hollow roads in LiDAR data from the Netherlands

Wouter B. Verschoof-van der Vaart[a,b,*,1], Juergen Landauer[c,1]

[a] Faculty of Archaeology, Leiden University, P.O. Box 9514, 2300 RA Leiden, The Netherlands
[b] Data Science Research Programme (Leiden Centre of Data Science), Leiden University, P.O. Box 9505, 2300 RA Leiden, The Netherlands
[c] Landauer Research, Wimpfener St. 10, 71642 Ludwigsburg, Germany

INFO ARTICLE

Historique de l'article :
Reçu le 22 juin 2020
Accepté le 18 octobre 2020
Disponible sur Internet le xxx

Keywords :
LiDAR
CNN
Machine learning
Archaeology
Hollow roads

ABSTRACT

The systematic mapping of hollow roads, traces of (post)medieval sunken cart tracks ways, can provide information on past human movement and historical route networks. However, the sheer amount of traces and of available high-quality data necessitates the use of computational methods for the automatic detection of these archaeological objects. Therefore, a novel approach, named CarcassonNet, has been developed that uses a combination of a Deep Learning convolutional neural network and image processing algorithms to detect and trace hollow roads in LiDAR data from the Netherlands. CarcassonNet has been specifically developed for the archaeological domain, focusing on being computationally light and suited for reconstructing partially preserved and intersected hollow roads. Instead of using the whole roads as input for the convolutional neural network, in CarcassonNet individual sections are used. This makes it much more cost-effective to create a sufficient training dataset, and makes the classification task (performed by the neural network) relatively simple, with better detection results. The output of CarcassonNet consists of two types of geospatial vectors that offer the opportunity to efficiently study the roads themselves and their precise location in the landscape (polygons), and the course of the roads and the resulting route network (lines). An experimental evaluation shows that CarcassonNet is able to effectively detect hollow roads, with a MCC score of 0.47. Furthermore, it is shown that using the Digital Terrain Model, instead of visualized LiDAR data (hillshade) improves the performance of the convolutional neural network. The results of this research offer opportunities to reconstruct vanished and abandoned (post)medieval routes and answer questions about human-landscape interactions.

1. Introduction

The ever-growing set of remotely sensed data offers opportunities to study the past on a landscape scale [1–3]. Especially, the advancement of Light Detection And Ranging (LiDAR) techniques [4–6] has facilitated the survey of archaeological traces that were hitherto difficult to investigate, due to forest and other vegetation cover [2,7]. For instance, hollow roads—traces of sunken cart track ways [8] caused by the repeated use of the same route—are hard to discern in the present-day landscape, but appear as clear longitudinal objects in LiDAR data (Fig. 1). While the manual analysis of remotely sensed data is a widespread practice in present-day archaeology [9], the sheer amount of available high-quality data necessitates the use of computer-aided methods for the (semi-)automatic detection of archaeological objects [10,11]. Recent years have seen a steady increase in the implementation of archaeological object detection techniques (see Reference [12] for an overview), with a trend towards Machine Learning and Deep Learning methods [13,14]. Implementations of Deep Learning in archaeology have predominantly used Convolutional Neural Networks (CNNs), a type of image feature extractor and classifier loosely inspired by the animal visual cortex [15]. These algorithms learn, comparable to other Machine Learning methods, to generalize from given examples rather than relying on human-defined parameters or rules. CNNs can be pre-trained on a large, generic dataset, and subsequently transfer learned or fine-tuned [16] on a relatively small, specific dataset. This is a clear advantage for domains where large, labelled datasets are scarce or non-existent, such as archaeology.

In archaeology, the detection rates of Deep Learning approaches are now reaching or even exceeding 80% for compact, localized objects, such as barrows and charcoal kilns [19–21]. However, the automated detection of more complex, large-scale landscape

* Corresponding author at: Faculty of Archaeology, Leiden University, P.O. Box 9514, 2300 RA Leiden, The Netherlands.
Adresses e-mail : w.b.verschoof@arch.leidenuniv.nl (W.B. Verschoof-van der Vaart), juergenlandauer@gmx.de (J. Landauer).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.culher.2020.10.009

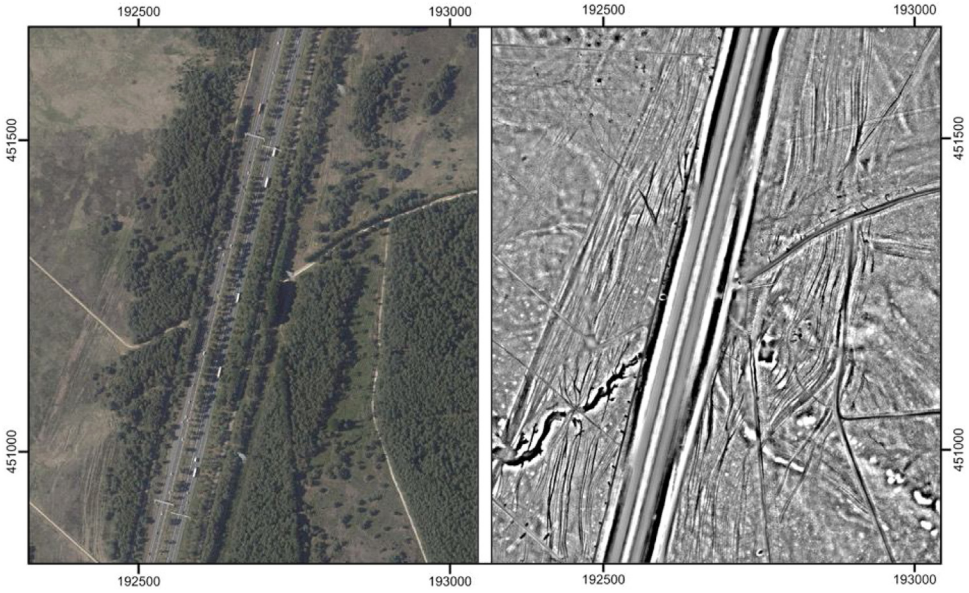

Fig. 1. An example of hollow roads on a recent aerial photograph (left) and on LiDAR data (right), visualized with Local Relief Model [17], from the Veluwe in the Netherlands (source: Reference [18], coordinates in Amersfoort/RD New, EPSG: 28992).

patterns, such as historical field systems and roads, has seen little application in the archaeological domain [22]. The few studies focussing on these patterns have generally used more traditional spatial and predictive modelling [23], or specialized (hand-crafted) feature extraction and analysis techniques to detect field systems [22,24] and roads [8,25–27]. Nevertheless, the systematic mapping of hollow roads can provide information on past human movement and historical route networks on both an inter-site and inter-regional level [28]. It can offer a better understanding of human-landscape interactions [29,30], as routes both reflect and influence (large-scale) cultural and landscape processes [27], and play a crucial role in the exchange of resources, knowledge, and ideas [31]. Especially the mapping of regional and local roads can supplement and expand the knowledge gained from historical written sources and cartographic data [32]. Therefore, in the study of historic routes we can distinguish two relevant levels of information: (1) the roads themselves and their precise location in the landscape; and (2) the course of the roads and the resulting route network [30]. These require different types of output from a detection model (see Section 2.2.4). While Deep Learning has successfully been implemented for modern road detection [33], these methods do not translate well to the problem at hand due to differences between hollow and modern roads: (1) hollow roads lack the uniformity that modern roads possess; (2) due to the long temporal and repeated use of specific route zones [34], hollow roads tend to manifest in the shape of several linear tracks with slightly different orientation and with multiple overlaps between the individual tracks; (3) due to geomorphological (e.g., erosion and drift) and anthropogenic interference (e.g., agricultural and building activities), hollow roads are often only partially preserved [35] and regularly are dissected by modern landscape objects. In this paper, we present a novel approach to automatically detect and trace hollow roads in LiDAR data, using a combination of Deep Learning and image processing algorithms. This approach, which we named *CarcassonNet*, has been specifically developed for the archaeological domain, focusing on being computationally light and suited for reconstructing partially preserved and intersected (hollow) roads. In the next Section (2), the research area and data is introduced, followed by an overview of CarcassonNet. In the subsequent Sections 3 and 4 the results of an experimental evaluation will be presented and discussed. The paper finishes with future developments planned (Section 5).

2. Materials and methods

2.1. Data & research area

In this research, three hollow road rich subareas (called Speuld, Terlet, and De Ginkel; in total 93.75 km^2) were investigated on the *Veluwe*, a region (circa 2200 km^2) in the central part of the Netherlands (Fig. 2). The Veluwe consists predominantly of forest and heathland, interspersed with agricultural fields and villages and towns of various size (for a detailed overview of the area see [12,20]). The Veluwe holds one of the densest concentrations of archaeological objects in the Netherlands, including prehistoric barrows [36], (post)medieval charcoal kilns [37], and (modern) traces of conflict [38]. Nowadays, the majority of these objects can be found under heath or forest cover. While this has almost certainly contributed to their preservation, this also hinders the physical investigation of these objects and restricts the survey of the surrounding landscape for potential new archaeological objects (see also Reference [7]). Previous research has shown that the Veluwe holds a legion of hollow roads [27], which thus far have gained little scientific attention. Unfortunately, no comprehensive overview of hollow roads on the Veluwe is available. Therefore, a reference standard [39] was manually created by the first author, who has ample experience in analysing LiDAR data and considerable knowledge of the archaeology of the research area (see also Section 2.2.1).

LiDAR data (Fig. 2, Table 1) of the area is freely available from the online repository PDOK [18] and the *Actueel Hoogtebestand Nederland* [41]. Hollow roads are distinguishable in LiDAR data as longitudinal objects (Fig. 1). These sunken lanes result from the intensive use of specific routes, as travellers preferred to keep to the beaten track to make for greater speed and reduce the risk of getting lost. Wherever the soil was soft and could be turned aside, a series of parallel lanes would develop. When a track became impassable due to seasonal conditions (e.g., weather) or general wear over time, travellers were forced to shift to an adjacent lane, resulting in the pattern of parallel cart-ruts [42], also called route zones [34]. It has been suggested that some of the hollow roads on the

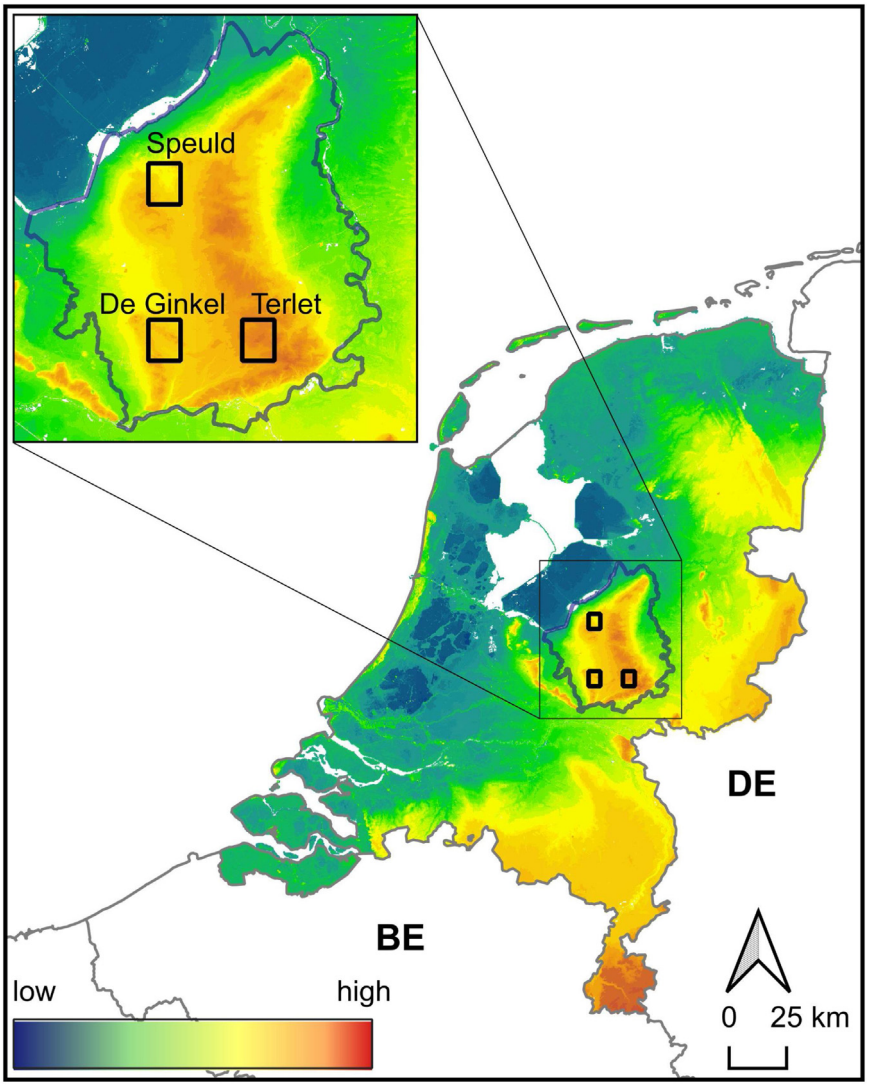

Fig. 2. The research area (outlined in blue) on a height model of the Netherlands showing the location of the three subareas used in black (source of the height model: [18]; coordinates in Amersfoort/RD New, EPSG: 28,992; amended from Reference [12]).

Table 1
The parameters of the LiDAR data used in this research, after Reference [40].

Meta-information LiDAR data	
Purpose	Water Management
Time of data acquisition	April 2010
Equipment	RIEGL LMS-Q680i Full-Waveform
Scan angle (whole FOV)	45°
Flying height above ground	600 m
Speed of aircraft (TAS)	36 m/s
Laser pulse rate	100,000 Hz
Scan rate	66 Hz
Strip adjustment	Yes
Filtering	Yes
Interpolation method	Moving planes
Point-density (pt per sq m)	6–10
DTM-resolution	0.5 m

Veluwe date back as far as the Bronze Age (2000–800 BC; [43]), although the earliest confirmed routes date to the late Middle Ages (1250–1500 AD [27]).

2.2. CarcassonNet

For the approach presented in this paper inspiration was taken from the well-known board game *Carcassonne* by Klaus-Jürgen Wrede [44], hence the name CarcassonNet. In this game, a medieval landscape, including roads, is constructed by laying small, square-shaped cards, so-called terrain tiles, end-to-end. While the roads depicted on these tiles differ little individually, extensive road networks of varying shapes can be constructed. The same applies for hollow roads: while these differ when examined as whole objects, the variation between small sections of hollow roads is minimal. Therefore, in CarcassonNet individual sections, as opposed to the whole hollow roads, are used as input images (comparable to the approach taken for the detection of Celtic fields in Reference [20] and for forest roads in Reference [45]). This has several advantages: (1) the use of sections (and therefore multiple pieces per single hollow road) makes it much more cost-effective to create a sufficient training dataset for Deep Learning algorithms; and (2) by using this method, a CNN is only used for image classification, a relatively simple task, instead of more complicated tasks such as segmentation. Another benefit of using classification over segmentation is that hollow roads are difficult to define on a pixel level because of their heterogeneous nature, especially if these traces have been 'transformed' by various natural and anthropogenic processes over time (see also Reference [46]). Therefore, using a CNN for classification will produce better detection results at a lower cost/effort [47]. CarcassonNet consists of three steps: preprocessing, classification,

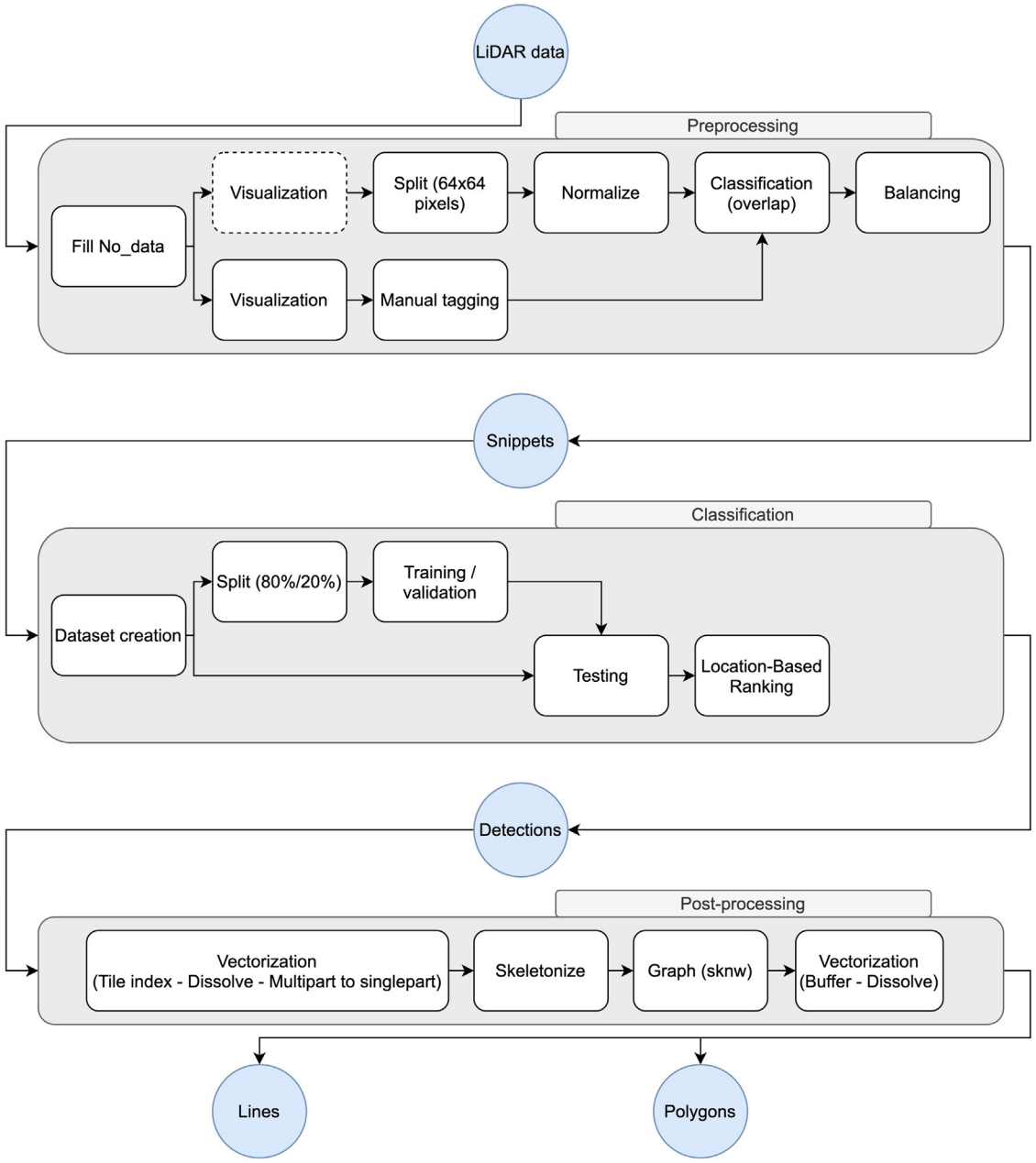

Fig. 3. Overview of the CarcassonNet workflow.

and post-processing (Fig. 3). In the following these steps will be discussed in detail.

2.2.1. Preprocessing

The LiDAR data was loaded into *QGIS 3.4 Madeira* [48] and processed with the *Fill_nodata* processing tool to reduce the number of no-data points. For the manual labelling, the data was visualized with the Local Relief Model [17] and Openness [49] visualizations from the *Relief Visualisation Toolbox 1.3* [5], based on the suitability of these visualizations to represent hollow roads. All hollow roads within the data were manually labelled in QGIS (also see Section 2.3), resulting in a reference standard consisting of a geospatial vector file (shapefile) with polygons demarcating hollow roads.

In earlier research it was noted that using the Digital Terrain Model (DTM) instead of visualized LiDAR data resulted in improved performance of Deep Learning models [50]. Therefore, for the training of our CNN, two datasets were created. In one dataset the LiDAR data was visualized with the hillshade visualization from the *Relief Visualisation Toolbox 1.3* [5]. The second dataset consisted of DTM data. The performance of both models was evaluated (see Section 3.2). The data (in Float32 format) in both datasets was split into snippets of 64×64 pixels with a 32-pixel overlap to all sides. The size of the snippets is based on the smallest size that yielded useful results. Subsequently, the snippets were normalized by subtraction of the central pixel value so that each snippet has pixel (or grayscale) values between 0 and 255. In order to create a binary dataset, with 'hollow road' and 'empty' snippets, a percentage of overlap was computed between every individual snippet and the polygons in the reference standard. This resulted in a dataset containing circa 10,000 positive examples and approximately 383,000 negative examples (Table 2).

As only a tiny fraction of the landscape consists of hollow roads, clearly a large imbalance between 'hollow road' and 'empty' examples was to be expected. Such an imbalanced dataset can

Table 2
The amount and ratio of snippets classified as "hollow road" or "empty" in the original, unbalanced dataset and the balanced dataset (used in this research).

Classification	Original Dataset		Balanced Dataset	
	Amount	Ratio	Amount	Ratio
Hollow road	10,185	2.6%	16,000	49.5%
Empty	383,401	97.4%	16,353	50.5%
Total	**393,586**	**100%**	**32,353**	**100%**

exert a major impact on the classification capacity of Deep Learning models and can result in bias and low performance [51]. Various approaches have been proposed to address this problem (see Reference [52] for a recent overview). In this research data pruning [53] and down and upsampling [54] were used to 'balance' the dataset. Empirically we found that placing only snippets with an overlap of 95% or more with the mapped hollow roads in our reference standard in the 'hollow road' class and only those with an overlap of 5% or less into the 'empty' class provided the best results when training the CNN. Snippets in between these thresholds were pruned from the dataset in order to make the two classes more distinguishable. Subsequently, the majority class ('empty') was downsampled, i.e., a large random portion of the examples (circa 98%) was removed, and the minority class was upsampled, i.e., randomly duplicated, until both classes contained approximately the same number of instances. This resulted in a dataset of circa 32,000 snippets, with circa 16,000 snippets in both classes (see Table 2). 80% of the snippets (circa 25,900) in the balanced dataset were randomly assigned to the training dataset. The remaining 20% (circa 6470) formed the validation dataset, used to evaluate the quality of the CNN during training.

2.2.2. Classification

For the classification step a Resnet-34 CNN architecture [55], developed using the *Fast.ai* library [56] on top of *Facebook's PyTorch Artificial Intelligence (AI)* development framework [57], was used. The Fast.ai library was chosen because it incorporates many state-of-the-art tuning methods to improve the performance of CNNs [58].

The CNN was pre-trained on ImageNet [59] and transfer learned [16] on our own balanced dataset (see above). Transfer learning was done following the 'progressive resizing' training scheme recommended in Fast.ai [56], using three separate phases. In the first phase, the CNN was trained for five epochs, during which the snippets in the training dataset retained their original size (64 by 64 pixels). In the second phase, the CNN was trained for four epochs but, contrary to the prior phase, the snippets in the training dataset were upscaled to 128 by 128 pixels. In both training phases all layers of the CNN were frozen (untrainable), except for the last fully connected layers. In the final phase, all layers of the CNN were unfrozen and the model was trained for three epochs on the upscaled snippets of 128 by 128 pixels. During training, the Adam optimization algorithm was used [13]. Discriminative layer training [60] and data augmentation (horizontal and vertical flipping [13]) were implemented to further improve the performance.

2.2.3. Location-Based Ranking

During the development of CarcassonNet the main sources of false positives were analysed. This showed that 'objects of confusion', with morphology comparable to hollow roads, generally caused these incorrect detections. Therefore, Location-Based Ranking (LBR; [21]) was implemented, to reduce these false positives and improve performance. LBR involves determining, ranking, and mapping of (present-day) landscape characteristics, such as subsoil and land-use, which have had an impact on the preservation and/or visibility of archaeology and result in objects of confusion

Table 3
The two ranks used in the Location-Based Ranking map for hollow roads on the Veluwe. Rank 1 has high preservation/visibility for hollow roads, while rank 2 has low preservation/visibility.

Rank	Landscape Feature	Source
2	*Plaggen soils*	Geological map (*Bodemkaart van Nederland*, scale 1:50.000)
	Modern agricultural areas	Digital Agricultural map (*Agrarisch Areaal Nederland*)
	Disturbed areas	Digital topographical map of the Netherlands (*Basisregistratie Grootschalige Topografie*)
	Built-up areas	Digital topographical map of the Netherlands (*Basisregistratie Grootschalige Topografie*)
	Roads	National road dataset (*Nationaal Wegen Bestand*).
1	Other	–

Table 4
Results of the validity test of the Location-Based Ranking map for hollow roads on the Veluwe.

Rank	Hollow Roads	
	m^2	Ratio
2	33,560	1.5%
1	2,237,170	98.5%
Total	2,270,730	100%

(for a detailed overview see Reference [21]). To determine the principal landscape characteristics involved, a broad-brush landscape characterization [1] was performed. This evaluation identified five processes as being the most detrimental to the preservation and visibility of hollow roads on the Veluwe and causing the most false positives: (post)medieval agricultural areas (*plaggen soils*; [61]), modern agricultural fields, disturbed areas (quarries, sod cutting, etc.) and to a lesser extent urbanized or built-up areas and modern paths and roads. The best chance for survival of hollow roads can be found in heathland. Based on this a ranked map of the research area was created, on which the assigned ranks correspond to the potential for the occurrence of hollow roads within that zone. Subsequently, hollow road detections were compared to this map and assigned different ranks. Detections in high-ranking zones (rank 1) are more likely to be hollow roads, while detections in low-ranking zones (rank 2) have a much higher likelihood of being false positives. Therefore, LBR can be used to reduce the amount of false positives by ignoring detections in low-ranking zones.

Based on the above, a simple two-tiered ranking map (Table 3 and Fig. 4) for the research area was developed, using open-source geo(morph)ological and topographical data from the online spatial data repository PDOK [18].

To test the validity of the proposed LBR map for the research area, the locations of all extant hollow roads in the Speuld subarea (see Fig. 4) were ranked. Table 4 shows the results of the ranking: more than 98.5% of the area covered with hollow roads can be assigned to the top rank (1). This shows the effectiveness of the ranking system and the landscape characteristics chosen. Furthermore, it demonstrates that by only considering rank 1 detections—ignoring detections in rank 2—the number of missed hollow roads will be low, while the number of false positives, caused by these zones, will be reduced.

2.2.4. Post-processing

To adequately investigate the location of the roads themselves, the course of the roads, and the resulting route network (see Section 1) requires different types of output of the post-processing step of CarcassonNet. For the location, areas demarcating the actual roads are sufficient. These areas could be used as a proxy for the road

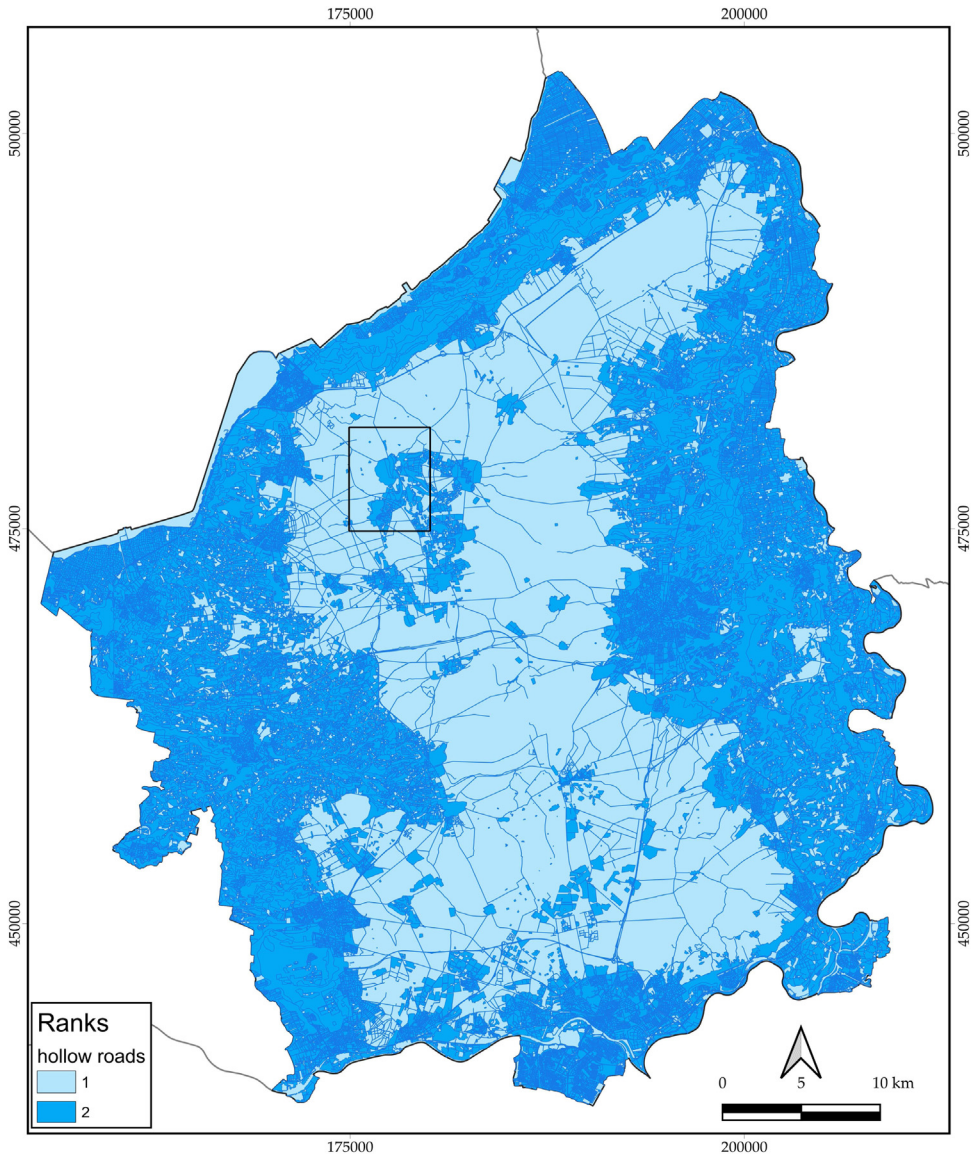

Fig. 4. The Location-Based Ranking map for hollow roads on the Veluwe, showing the Speuld subarea used for validation in black (coordinates in Amersfoort/RD New, EPSG: 28992).

courses, but to increase the usability of the results for the interpretation of the wider route network, these polygons should be converted into lines.

Therefore, in the post-processing step the results of the preceding classification step, i.e., snippets with a class label and a confidence score, are converted into geospatial polygons and subsequently into lines, both usable in a GIS (Fig. 5). All snippets classified as roads are loaded into *QGIS 3.4 Madeira* [48]. These raster files are turned into vectors, combined into larger polygons, and turned from multipart to individual singlepart features with the *Tileindex*, *Dissolve*, and *Multipart_to_singlepart* processing tools. Subsequently, these polygons are turned into lines using the approach taken by Reference [62]. A 'skeleton' is created from the polygons with the scikit-image *Skeletonize* package and subsequently rendered into a graph structure with the *sknw* package in *Python 3.7*. The result is a set of closely spaced points depicting the course of the detected roads. These are turned into vectors with the *Buffer* and *Dissolve* processing tools in QGIS.

2.3. Feedback loop

In the training process of CarcassonNet, a feedback loop between the archaeological interpreter and the classification algorithm was implemented to enhance the quality of the (training) dataset and the classification model (Fig. 6). Initially, the first author rapidly labelled the hollow roads within the dataset (see Section 2.2.1). Subsequently, the algorithm was trained and tested on this data. The incorrect detections (false positives) were manually compared to the LiDAR data to determine their cause. This resulted into multiple unlabelled hollow roads, also called 'new positives' [63], that were missed during the initial tagging. The training dataset was updated with these new positives and used to retrain the algorithm. This process was repeated two times till no new positives were detected. This approach allows for the rapid labelling of the initial training dataset, reducing the time needed for preprocessing. Furthermore, it results in the improved performance and a thorough training dataset.

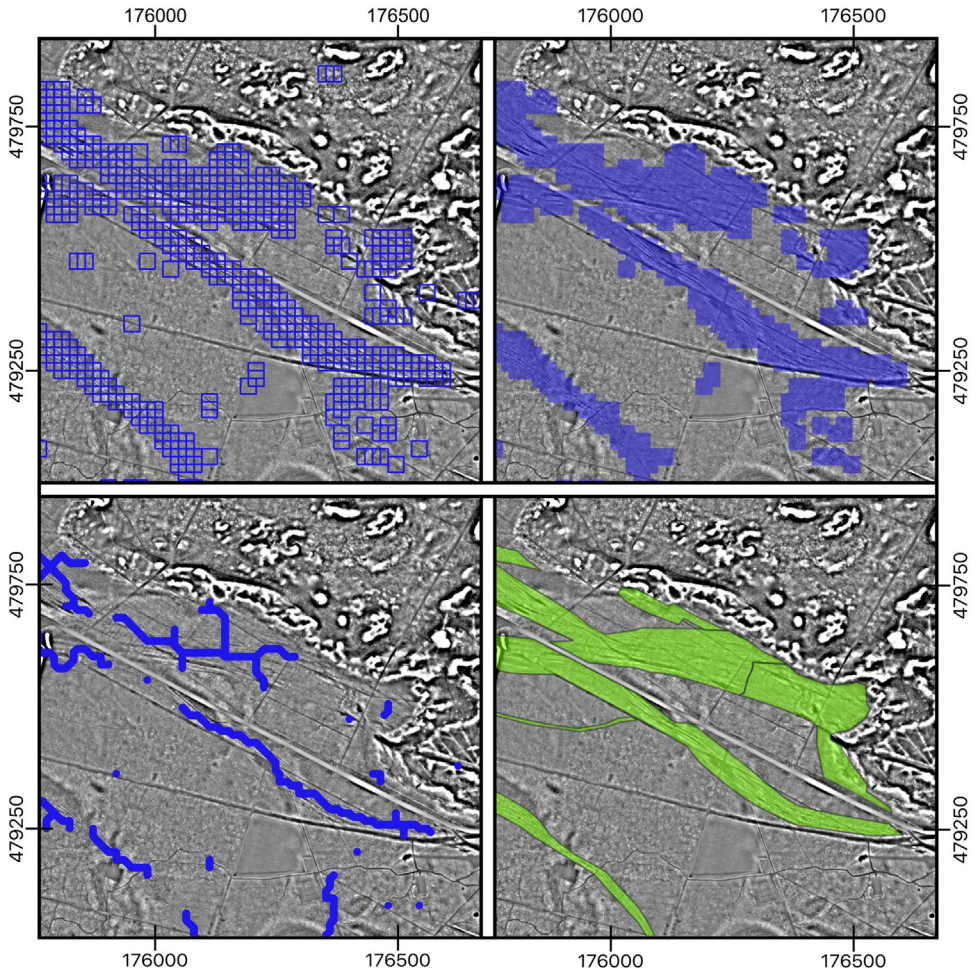

Fig. 5. Excerpts of LiDAR data, visualized with the Local Relief Model [17], showing: the results of the classification (top left); the derived polygons showing the location of the hollow roads (top right); the derived lines depicting the route network (bottom left); and the manually labelled reference standard in green (bottom right); source of the height model: [18]; coordinates in Amersfoort/RD New, EPSG: 28992).

3. Results

3.1. Metrics

To evaluate the workflow, the areas (in m^2) of false negatives (FN), false positives (FP), true negatives (TN), and true positives (TP) were determined, following the approach taken for Celtic fields in Reference [21]. The following metrics were calculated: Precision (P), Recall (R), F1-Score (F1), Accuracy (ACC), and Matthews Correlation Coefficient (MCC; Equations 1–5). Precision measures how many of the selected items are relevant. Recall gives a measure of how many relevant objects are selected. The F1-score is the harmonic average of the precision and recall and a measure of the model's performance [64]. Accuracy gives the ratio of correctly predicted instances to the total of instances in the dataset. MCC is a measure of the correlation between the observed and predicted binary (two-class) classification, even if the classes are very imbalanced [65]. Therefore, MCC is a more reliable indicator of quality of the model over the other metrics used [51]. Precision, recall, F1-score, and accuracy are restricted between 0 and 1, while MCC is bound between −1 and 1. For all metrics higher values indicate a better performance. We are aware that the above metrics, while adequate to evaluate classification tasks, are suboptimal to access the quality of connected road networks [62,66]. However, the above metrics are more established in remote sensing and archaeological Machine Learning research and are therefore easier to assess and compare with work from other authors (see also Reference [20]).

Equations 1–5: Precision (1), recall (2), F1-score (3), accuracy (4) and MCC (5).

$$Precision = TP/(TP + FP)$$

$$Recall = TP/(TP + FN)$$

$$F1 - score = 2 * (Recall * Precision)/(Recall + Precision)$$

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)$$

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

3.2. Experimental evaluation

In our experiments, the ResNet-34 CNN was trained as detailed above on the balanced dataset. Subsequently, the trained CNN was evaluated on the original dataset from the Veluwe (see Table 2). This dataset was used as an unbalanced dataset better represents the real-world situation of scarce archaeological objects in a complex landscape (see Reference [21]). All experiments were performed on

Table 5
the results of the experimental evaluation showing the difference in performance between visualized data (HS) and digital terrain model data (DTM) and the use of Location-Based Ranking (LBR).

Data	LBR	TP (m^2)	FP (m^2)	TN (m^2)	FN (m^2)	R	P	F1	ACC	MCC
HS	no	3,901,040	6,931,200	89,152,100	3,731,300	0.511	0.360	0.423	0.897	0.375
DTM	no	4,997,293	8,300,800	87,782,500	2,635,047	**0.655**	0.376	0.478	0.895	0.444
HS	yes	3,882,350	6,214,600	89,868,700	3,748,990	0.509	0.385	0.438	**0.904**	0.391
DTM	yes	4,967,470	7,243,900	88,839,400	2,664,868	0.651	**0.407**	**0.501**	**0.904**	**0.466**

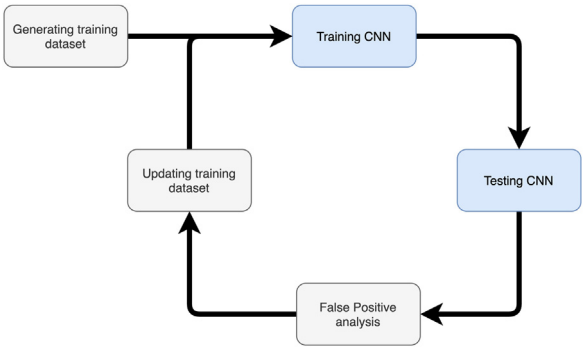

Fig. 6. Schematic representation of the feedback loop used in this research with the steps done by the researcher in white and the steps done by the CNN in blue.

a NVIDIA GTX 1060 GPU. During the experimental evaluation the confidence threshold was empirically set to 0.80. The results of the experimental evaluation are shown in Table 5. CarcassonNet has a recall (R) of 0.51 and a precision (P) of 0.36. It reaches a F1-score of 0.42, an accuracy of 0.89, and a MCC of 0.38 on the dataset with hillshade visualization (Table 5). When using DTM data (instead of the hillshade) the performance increases with circa 6 points to an F1-score of 0.48 and a MCC of 0.44. Implementing LBR further increases the performance with circa 1-2 points to a MCC of 0.39 for hillshade and 0.47 for DTM. Table 5 clearly shows the problem of using the accuracy and F1-score metrics on an imbalanced dataset. The accuracy is very high, due to a bias towards the majority category (in this case TN). On the other hand, the recall, precision, and F1-score are relatively low, because these metrics are dependent on which class is defined as positive (and ignores the TN category). Therefore, the only metric that gives a reliable result is MCC, as it takes into account the balance ratios of all four categories (TP, FN, TN, and FP; [67]).

Visual analyses of the results show that in general larger or wider bundles of multiple hollow roads are well recognized by the model (Fig. 7). In contrast, smaller bundles or individual hollow roads that are significantly narrower than the snippet size (64 by 64 pixels) are often missed or only partially detected. This might seem related to the overlap threshold used in the preprocessing step (see Section 2.2.1). However, lowering the overlap threshold to 25, 50, or 75% only resulted in a decline in performance of the model and no improvement in the prediction quality of 'thinner' roads. Therefore, the overlap threshold seems to be of no influence to the detection of smaller bundles or individual hollow roads. Conceivably, the balance of the classes in the dataset might be of influence to the performance and accuracy. False positives are generally caused by objects of confusion, which are morphologically comparable to hollow roads (Fig. 8). There seems to be no clear pattern in these, as many of the common objects of confusion are removed through LBR. Most of the false positives consist of individual or a few detections (or snippets) clustered together, whereas most true positives consist of a larger amount of clustered detections.

The results of the post-processing step of CarcassonNet (Figs. 5, 7, and 9) show that the conversion of the detections into polygons and lines works adequately on individual or smaller bundles of hollow roads. However, when a polygon of a large bundle of hollow roads (with many roads running parallel) is converted to lines a problem arises in which branches, perpendicular on the course of the road, are generated (see Fig. 9). This reduces the readability of the road network. Further research is needed in order to resolve this problem.

Table 6
The results of the implementation of the feedback loop per subarea.

Subarea	Original (m^2)	Updated (m^2)	Difference (m^2)	Difference (%)
Speuld	2,165,460	2,270,730	105,270	4.86
Terlet	2,687,350	2,849,630	162,280	6.04
De Ginkel	2,382,310	2,512,030	129,720	5.45
Total	**7,235,120**	**7,632,390**	**397,270**	**5.49**

3.3. Evaluation of the feedback loop

During the development of CarcassonNet, a feedback loop was used to detect new positives and update the training dataset (see Section 3.2). It was assumed that this approach would improve the quality of the dataset and the performance of the CNN model. In Table 6 the total area of the manually labelled road polygons in the original and updated datasets is shown per subarea (see Fig. 2). The difference between the datasets, circa 397,000 m^2, equals the total area of new positives that were identified. This shows that in the initial labelling on average 5.5% of the hollow roads in the Veluwe dataset were missed. Furthermore, two iterations of the feedback loop improved the performance (MCC) of CarcassonNet by circa 0.2. These results show the benefit of using the feedback loop in the development of datasets for automated detection methods.

4. Discussion

4.1. CarcassonNet versus modern road detection approaches

The results of the experimental evaluation (Table 5) show that CarcassonNet is able to detect hollow roads in LiDAR data. The strategies implemented in this research resulted in a performance (MCC) of 0.47. Validating the performance of CarcassonNet by comparing it with other automated historical road detection methods proves difficult, as the few studies do not provide (comparable) metrics [26,27,68]. Vletter estimates a precision of over 0.80 for the approach presented in Reference [68] (i.e., over 80% of the detections are segments of roads or paths). However, no other metrics are given. The performance (F1) of recent Machine Learning and Deep Learning detection methods for modern roads in remotely sensed data varies widely, between 0.67 and 0.96 (Table 7). Our approach (F1-score of 0.50) seems to be lower. The main reason for this is the difference between hollow roads and modern roads, as hollow roads consist of dissected, re-used, overlapping tracks that lack uniformity (see Section 1). These factors make it much more complicated to adequately detect hollow roads, compared to modern roads. Furthermore, as shown above the F1-score is not an adequate metric to evaluate the performance of binary road classification. Unfortunately, it has been impossible to calculate the MCC

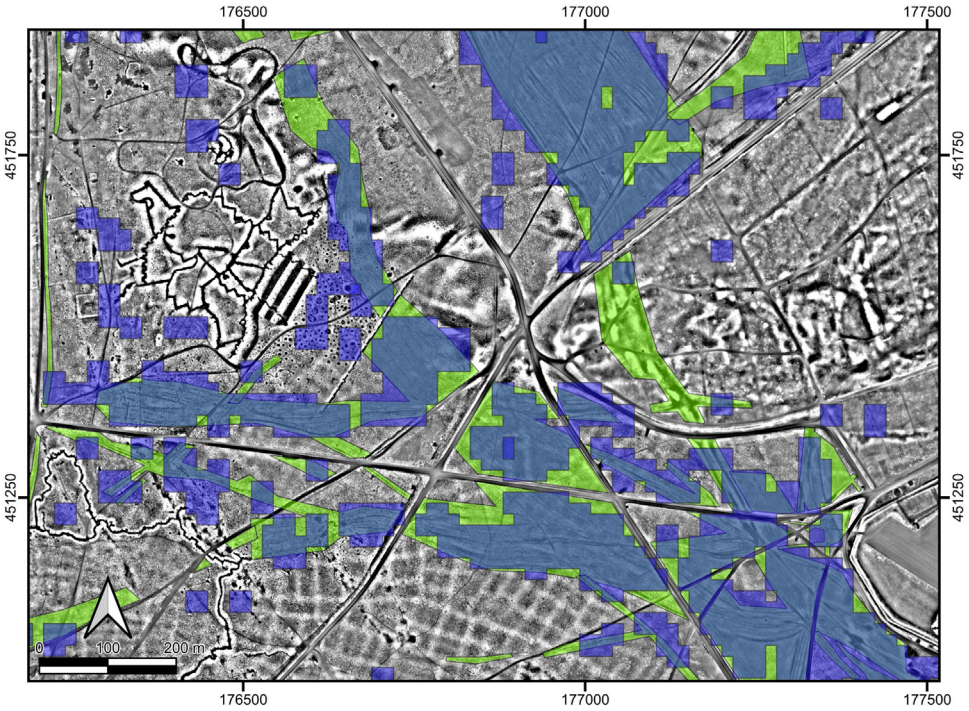

Fig. 7. Excerpt of LiDAR data from the De Ginkel subarea, visualized with Local Relief Model [17], showing the areas classified as hollow road by CarcassonNet in blue and the reference standard in green (source of the height model: [18]; coordinates in Amersfoort/RD New, EPSG: 28992).

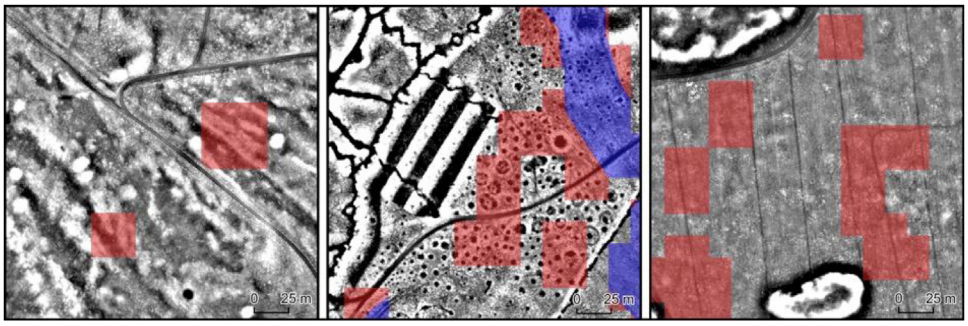

Fig. 8. Excerpts of LiDAR data, visualized with the Local Relief Model [17], showing some examples of false positives (red) resulting from objects of confusion in the research area: geological objects (left); other archaeological objects (centre); objects related to (modern) forestry (right; source of the height model: [18]).

Table 7
Overview of recent Machine Learning and Deep Learning (modern) road detection methods.

Research	Type of roads	Type of Data	Method	F1-score
CarcassonNet	Hollow roads	LiDAR	Deep Learning	0.501
Ferraz et al. [45]	Forest roads	LiDAR	Machine Learning	0.849
Li et al. [69]	Roads (Shaoshan City)	Satellite imagery	Machine Learning	0.839
Nachmany & Alemohammed [70]	Roads (Khartoum)	Satellite imagery	Deep Learning	0.770
Pradhan and Ibrahim Sameen [71]	Roads (highways)	LiDAR/Orthophotos	Machine Learning	0.808
Xu et al. [72]	Roads	Satellite imagery	Deep Learning	0.957
Van Etten et al. [66]	Roads (Las Vegas, Paris, Shanghai, Khartoum)	Satellite imagery	Deep Learning	0.666

for the below methods, as no exact information on the amount of TP, FN, TN, and FP was given.

4.2. DTM versus visualized LiDAR data

The results show that the performance of CarcassonNet varies between visualized and DTM data (Table 5). This is most likely related to information loss, as proposed by Shannon's channel coding theorem [73]. This theorem from Information Theory states that each processing of information from one representation into another (in our case from DTM to visualized LiDAR data) reduces, for instance due to rounding errors, but never improves the information content. Therefore, the DTM will contain more informational, less noisy data, as compared to the visualized data.

4.3. Archaeological implementation

The method and results presented in this research can be used to get a better understanding of (post)medieval routes on the Veluwe. By combining our mapped (regional and local) roads with informa-

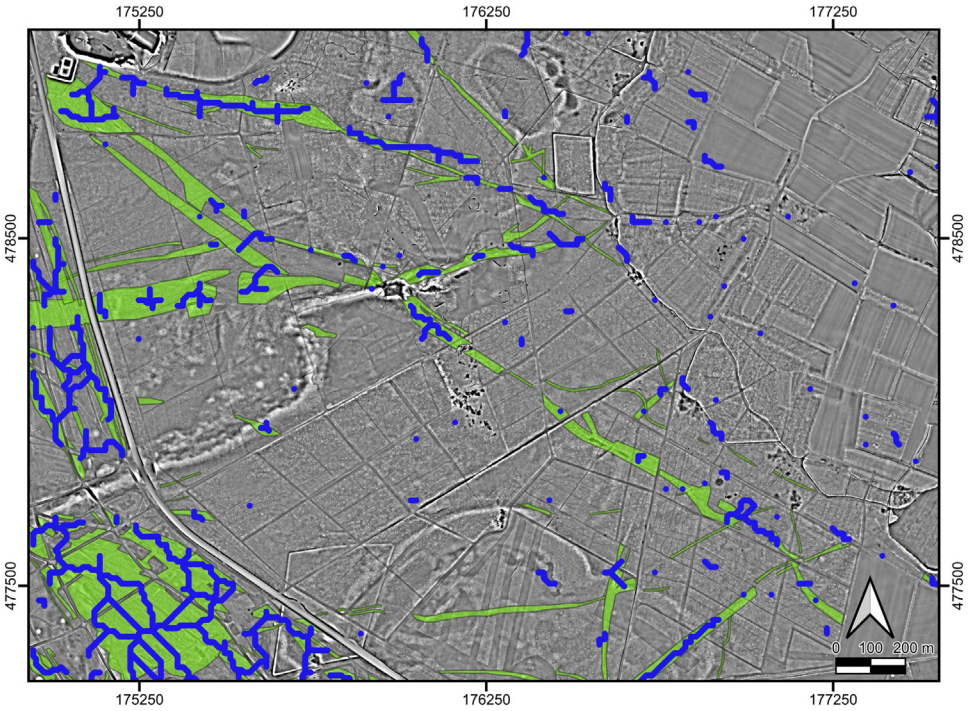

Fig. 9. Excerpt of LiDAR data from the Speuld subarea, visualized with Local Relief Model [17], showing the lines depicting the route network generated by CarcassonNet in blue and the reference standard in green (source of the height model: [18]; coordinates in Amersfoort/RD New, EPSG: 28992).

tion gained from historical written sources, cartographic data, and modelling [27] a more complete picture of the route networks on the Veluwe can be gained. This will enable the investigation of the relative chronology of these roads (see Reference [74]). Furthermore, our method offers opportunities to efficiently investigate archaeological hypotheses on a landscape scale. For instance, the relation between prehistoric barrows (as possible road markers) and hollow roads [75] can be investigated. Also, the relation between hollow roads and forests, which according to Jager were avoided on the Veluwe due to active highwaymen [42], can be investigated.

5. Conclusions

In this research we presented a novel approach, named CarcassonNet, to automatically detect and trace hollow roads in LiDAR data, using a combination of Deep Learning CNNs and image processing algorithms. CarcassonNet consists of three steps: preprocessing, classification, and post-processing. Contrary to other research, CarcassonNet uses individual sections of roads as input, instead of whole roads. This makes it much more cost-effective to create a sufficient training dataset, and makes the classification task (performed by the neural network) relatively simple, leading to better detection results at a lower cost/effort. The output of CarcassonNet consists of two types of geospatial vectors to efficiently study the roads themselves and their precise location in the landscape (polygons), and the course of the roads and the resulting route network (lines). Experiments show that CarcassonNet is able to effectively detect hollow roads in data from the Veluwe, with a MCC score of 0.47. These experiments also show that using the Digital Terrain Model, instead of visualized LiDAR data (i.e., hillshade) improves the performance with circa 6 points. Implementing Location-Based Ranking, to reduce the number of false positives, further increases the performance with circa 2 points.

Further research will focus on the improvement of the performance of CarcassonNet on small, individual hollow roads, for instance by changing to a segmentation architecture, such as U-Net [76], or by implementing segmentation in the post-processing step. Also, further incorporation of CarcassonNet into QGIS is envisioned. Finally, the development of CarcassonNet is part of an on-going PhD research (in the Data Science Research Programme at the Faculty of Archaeology at Leiden University, the Netherlands) to explore the potential of Deep Learning for the mapping of archaeological objects in remotely sensed data. Earlier research focussed on detecting barrows, Celtic fields, and charcoal kilns, resulting in the WODAN workflow [20,21]. Future research will involve the use of both models in conjunction to detect archaeological objects in a case study area, and investigate the quantitative and qualitative knowledge gain of employing automated detection methods.

Funding

This work was in part supported by the Data Science Research Programme (Leiden University).

Acknowledgements

We would like to thank Karsten Lambers (Faculty of Archaeology, Leiden University) and Wojtek Kowalczyk (LIACS, Leiden University) for their valuable input during the writing of this article. We are also grateful to the editor and reviewers for their constructive comments.

Références

[1] D. Cowley, Remote sensing for archaeology and heritage management — site discovery, interpretation and registration, in: D.C. Cowley (Ed.), Remote Sensing for Archaeological Heritage Management. Proceedings of the 11th EAC Heritage Management Symposium, Reykjavîk, Iceland, 25–27 March 2010, Europae Archaeologia Consilium, Brussel, 2011, pp. 43–55.
[2] B.J. Devereux, G.S. Amable, P. Crow, A.D. Cliff, The potential of airborne lidar for detection of archaeological features under woodland canopies, Antiquity 79 (2005) 648–660, http://dx.doi.org/10.1017/S0003598X00114589.
[3] K. Lambers, Airborne and space borne remote sensing and digital image analysis in archaeology, in: C. Siart, M. Forbriger, O. Buben-

zer (Eds.), Digital Geoarchaeology: New Techniques for Interdisciplinary Human-Environmental Research, Springer, Cham, 2018, pp. 109–122, http://dx.doi.org/10.1007/978-3-319-25316-9_7.
[4] R.S. Opitz, An overview of airborne and terrestial laser scanning in archaeology, in: R.S. Opitz, D.C. Cowley (Eds.), Interpreting Archaeological Topography. Airbornce Laser Scanning, 3D Data and Ground Observation, Oxbow Books, Oxford and Oakville, 2013, pp. 13–31.
[5] Ž. Kokalj, R. Hesse, Airborne Laser Scanning Raster Data Visualization. A Guide to Good Practice, Založba ZRC, Ljubljana, 2017.
[6] S. Crutchley, P. Crow, Using Airborne Lidar in Archaeological Survey: The Light Fantastic, 2nd ed., Historic England, Swindon, 2018 https://historicengland.org.uk/images-books/publications/using-airborne-lidar-in-archaeological-survey/,. (Accessed 28 June 2017).
[7] H. Kenzler, K. Lambers, Challenges and perspectives of woodland archaeology across Europe, in: F. Gilingny, F. Djindjian, L. Costa, P. Moscati, S. Robert (Eds.), CAA2014: 21st Century Archaeology, Concepts, Methods and Tools. Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Archaeopress, 2015, pp. 73–80.
[8] A. Kirchner, N. Herrmann, R. Stadtmann, T. Lahmer, L.-M. Hille, T. Steinbrecher, M. Sauerwein, Spatial analysis of hollow ways in the Hildesheimer Wald Mountains (Lower Saxony, Germany) as a model for mountainous regions of Central Europe, Erdkunde 74 (2020) 1–14, http://dx.doi.org/10.3112/erdkunde.2020.01.01.
[9] D.C. Cowley, In with the new, out with the old? Auto-extraction for remote sensing archaeology, Proceedings of SPIE 8532 (2012), 853206, http://dx.doi.org/10.1117/12.981758.
[10] R. Bennett, D. Cowley, V. De Laet, The data explosion: tackling the taboo of automatic feature recognition in airborne survey data, Antiquity 88 (2014) 896–905, http://dx.doi.org/10.1017/S0003598X00050766.
[11] A. Bevan, The data deluge, Antiquity 89 (2015) 1473–1484, http://dx.doi.org/10.15184/aqy.2015.102.
[12] K. Lambers, W.B. Verschoof-van der Vaart, Q.P.J. Bourgeois, Integrating remote sensing, machine learning, and citizen science in Dutch archaeological prospection, Remote Sensing 11 (2019) 794, http://dx.doi.org/10.3390/rs11070794.
[13] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, The MIT Press, Cambridge, MA, 2016.
[14] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444, http://dx.doi.org/10.1038/nature14539.
[15] J.E. Ball, D.T. Anderson, C.S. Chan, Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community, Journal of Applications in Remote Sensing 11 (2017), 042609, http://dx.doi.org/10.1117/1.JRS.11.042609.
[16] A.S. Razavian, H. Azizpour, J. Sullivan, S. Carlsson, CNN features off-the-shelf: an astounding baseline for recognition, in: IEEE CVPR Workshop, 2014, pp. 806–813, http://dx.doi.org/10.1109/CVPRW.2014.131.
[17] R. Hesse, LiDAR-derived local relief models-a new tool for archaeological prospection, Archaeological Prospection 17 (2010) 67–72, http://dx.doi.org/10.1002/arp.374.
[18] Publieke Dienstverlening Op de Kaart (PDOK), (n.d.). https://www.pdok.nl/. (Accessed 12 June 2020).
[19] Ø.D. Trier, A.-B. Salberg, L.H. Pilø, Semi automatic mapping of charcoal kilns from airborne laser scanning data using deep learning, in: M. Matsumoto, E. Uleberg (Eds.), CAA2016: Oceans of Data. Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology, Archaeopress, Oxford, 2018, pp. 219–231.
[20] W.B. Verschoof-van der Vaart, K. Lambers, 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 (2019) 31–40, http://dx.doi.org/10.5334/jcaa.32.
[21] W.B. Verschoof-van der Vaart, K. Lambers, W. Kowalczyk, Q.P.J. Bourgeois, Combining deep learning and location-based ranking for large-scale archaeological prospection of LiDAR data from the Netherlands, ISPRS International Journal on Geo Information 9 (2020) 293, http://dx.doi.org/10.3390/ijgi9050293.
[22] A. Traviglia, A. Torsello, Landscape pattern detection in archaeological remote sensing, Geosciences 7 (2017) 128, http://dx.doi.org/10.3390/geosciences7040128.
[23] P. Verhagen, L. Nuninger, M.R. Groenhuijzen, Modelling of pathways and movement networks in archaeology: an overview of current approaches, in: P. Verhagen, J. Joyce, M.R. Groenhuijzen (Eds.), Finding the Limits of the Limes: Modelling Demography, Springer, Cham, 2019, http://dx.doi.org/10.1007/978-3-030-04576-0_11.
[24] B. Figorito, E. Tarantino, Semi-automatic detection of linear archaeological traces from orthorectified aerial images, International Journal of Applications in Earth Observation and Geoinformation 26 (2014) 458–463, http://dx.doi.org/10.1016/j.jag.2013.04.005.
[25] C. Sevara, M. Pregesbauer, M. Doneus, G. Verhoeven, I. Trinks, Pixel versus object — a comparison of strategies for the semi-automated mapping of archaeological features using airborne laser scanning data, Journal of Archaeological Science: Reports 5 (2016) 485–498, http://dx.doi.org/10.1016/j.jasrep.2015.12.023.
[26] W.F. Vletter, (Semi) automatic extraction from airborne laser scan data of roads and paths in forested areas, in: D.G. Hadjimitsis, K. Themistocleous, S. Michaelides, G. Papadavid (Eds.), Second Int. Conf. Remote Sens. Geoinf. Environ., 2014.
[27] W.F. Vletter, R.J. van Lanen, Finding vanished routes: applying a multi-modelling approach on lost route and path networks in the Veluwe Region, The Netherlands, Rural Landscapes: Society, Environment, History 5 (2018) 1–19, http://dx.doi.org/10.16993/rl.35.
[28] T.J. Wilkinson, C. French, J.A. Ur, M. Semple, The geoarchaeology of route systems in northern Syria, Geoarchaeology 25 (2010) 745–771, http://dx.doi.org/10.1002/gea.20331.
[29] D. Mlekuž, Roads to nowhere? Disentangling meshworks of holloways, in: Z. Czajlik, A. Bödo″cs (Eds.), Archaeology and Remote Sensing from the Baltic to the Adriatic, Institute of Archaeological Sciences, Faculty of Humanities, Eötvös Loránd University, Budapest, 2013, pp. 37–42.
[30] L. Nuninger, R. Opitz, P. Verhagen, T. Libourel, C. Laplaige, S. Leturcq, N. Le Voguer, C. Fruchart, Ž. Kokalj, X. Rodier, Developing FAIR ontological pathways: linking evidence of movement in Lidar to models of human behaviour, Journal of Computer Applications in Archaeology 3 (2020) 63–75, http://dx.doi.org/10.5334/jcaa.46.
[31] M. Løvschal, Ways of wandering — in the late bronze age barrow landscape of the himmerland-area, Denmark, in: D. Fontijn, A.J. Louwen, S. Van Der Vaart, K. Wentink (Eds.), Beyond Barrows. Current research on the structuration and perception of the Prehistoric Landscape through Monuments, Sidestone Press, Leiden, 2013, pp. 225–250.
[32] M. Slamova, B. Belacek, J. Beljak, N. Pazinova, F. Chudy, Dependence of the medieval settlements and historical roads to the natural environment around the deserted castle in Zvolen (Slovakia), Procedia - Social and Behavioral Sciences 120 (2014) 213–223, http://dx.doi.org/10.1016/j.sbspro.2014.02.098.
[33] A. Abdollahi, B. Pradhan, N. Shukla, S. Chakraborty, A. Alamri, Deep learning approaches applied to remote sensing datasets for road extraction: a state-of-the-art review, Remote Sensing 12 (2020) 1444, http://dx.doi.org/10.3390/rs12091444.
[34] R.J. van Lanen, M.C. Kosian, B.J. Groenewoudt, T. Spek, E. Jansma, Best travel options: modelling Roman and early-medieval routes in the Netherlands using a multi-proxy approach, Journal of Archaeological Science: Reports 3 (2015) 144–159, http://dx.doi.org/10.1016/j.jasrep.2015.05.024.
[35] V. De Laet, E. Paulissen, M. Waelkens, Methods for the extraction of archaeological features from very high-resolution Ikonos-2 remote sensing imagery, Hisar (southwest Turkey), Journal of Archaeological Science 34 (2007) 830–841, http://dx.doi.org/10.1016/j.jas.2006.09.013.
[36] Q.P.J. Bourgeois, Monuments on the Horizon. The Formation of the Barrow Landscape throughout the 3rd and 2nd Millennium BC, Sidestone Press, Leiden, 2013.
[37] B. Groenewoudt, Charcoal burning and landscape dynamics in the Early Medieval Netherlands, Ruralia 6 (2007) 327 https://www.brepolsonline.net/doi/pdf/10.1484/M.RURALIA-EB.3.1150,. (Accessed 14 May 2018).
[38] M. van der Schriek, W. Beex, The application of LiDAR-based DEMs on WWII conflict sites in the Netherlands, Journal of Conflict Archaeology 12 (2017) 94–114, http://dx.doi.org/10.1080/15740773.2017.1440960.
[39] C.D. Manning, P. Raghavan, H. Schütze, Introduction to Information Retrieval, Cambridge University Press, Cambridge, 2009.
[40] N. van der Zon, Kwaliteitsdocument AHN-2 https://assets.amsterdam.nl/publish/pages/704401/kwaliteitsdocumentahn.pdf, 2013,. (Accessed 1 August 2017).
[41] Actueel Hoogtebestand Nederland (AHN), (n.d.). http://ahn.arcgisonline.nl/ahnviewer/. (Accessed 12 June 2020).
[42] S.W. Jager, A prehistoric route and ancient cart-tracks in the Gemeente of Anloo (Province of Drenthe), Palaeohistoria 27 (1985) 185–245.
[43] J.A. Brongers, Air Photography and Celtic Field Research in the Netherlands, Rijksdienst voor het Oudheidkundig Bodemonderzoek, Amersfoort, 1976.
[44] Carcassonne, (n.d.). https://www.zmangames.com/en/products/carcassonne/. (Accessed 24 March 2020).
[45] A. Ferraz, C. Mallet, N. Chehata, Large-scale road detection in forested mountainous areas using airborne topographic lidar data, ISPRS Journal of Photogrammetry and Remote Sensing 112 (2016) 23–36, http://dx.doi.org/10.1016/j.isprsjprs.2015.12.002.
[46] A. Guyot, L. Hubert-Moy, T. Lorho, Detecting Neolithic burial mounds from LiDAR-derived elevation data using a multi-scale approach and machine learning techniques, Remote Sensing 10 (2018) 225, http://dx.doi.org/10.3390/rs10020225.
[47] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, M.S. Lew, Deep learning for visual understanding: a review, Neurocomputing 187 (2016) 27–48, http://dx.doi.org/10.1016/j.neucom.2015.09.116.
[48] QGIS Development Team, QGIS Geographic Information System http://qgis.org, 2017.
[49] R. Yokoyama, M. Shirasawa, R.J. Pike, Visualizing topography by openness: A new application of image processing to digital elevation models, Photogrammetric Engineering and Remote Sensing 68 (2002) 257–265.
[50] J. Landauer, R. Hesse, Machine learning for large area archaeological feature detection applying transfer learning to airborne lidar data, in: CHNT 24, 2019 – Abstr., 2019, pp. 1–3 https://www.chnt.at/wp-content/uploads/Machine-learning-for-large-area-archaeological-feature-detection.pdf,. (Accessed 12 June 2020).
[51] A. Luque, A. Carrasco, A. Martín, A. de las Heras, The impact of class imbalance in classification performance metrics based on the binary confusion matrix, Pattern Recognition 91 (2019) 216–231, http://dx.doi.org/10.1016/j.patcog.2019.02.023.
[52] J.M. Johnson, T.M. Khoshgoftaar, Survey on deep learning with class imbalance, Journal of Big Data 6 (2019) 27, http://dx.doi.org/10.1186/s40537-019-0192-5.
[53] A. Angelova, Y. Abu-Mostafa, P. Perona, Pruning training sets for learning of object categories, in: Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, IEEE, 2005, pp. 494–501.

[54] H. He, E.A. Garcia, Learning from imbalanced data, IEEE Transactions on Knowledge and Data Engineering 21 (2009) 1263–1284, http://dx.doi.org/10.1109/tkde.2008.239.
[55] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, NV, 2016, pp. 770–778.
[56] J. Howard, S. Gugger, fastai: A Layered API for Deep Learning https://arxiv.org/abs/2002.04688, 2020,. (Accessed 4 May 2020).
[57] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, A. Lerer, Automatic differentiation in PyTorch, in: NIPS 2017 Autodiff Workshop: The Future of Gradient-based Machine Learning Software and Techniques, NIPS, 2017.
[58] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, M. Li, Bag of Tricks for Image Classification With Convolutional Neural Networks https://arxiv.org/abs/1812.01187, 2018,. (Accessed 4 May 2020).
[59] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, L. Fei-Fei, ImageNet large scale visual recognition challenge, International Journal of Computer Vision 115 (2015) 211–252, http://dx.doi.org/10.1007/s11263-015-0816-y.
[60] A. Brock, T. Lim, J.M. Ritchie, N. Weston, FreezeOut: Accelerate Training by Progressively Freezing Layers http://arxiv.org/abs/1706.04983, 2017,. (Accessed 5 May 2020).
[61] H.P. Blume, P. Leinweber, Plaggen soils: landscape history, properties, and classification, Journal of Plant Nutrition and Soil Science 167 (2004) 319–327, http://dx.doi.org/10.1002/jpln.200420905.
[62] A. Van Etten, City-scale Road Extraction From Satellite Imagery http://arxiv.org/abs/1904.09901, 2019.
[63] M.F. Meyer, I. Pfeffer, C. Jürgens, Automated detection of field monuments in digital terrain models of Westphalia using OBIA, Geosciences 9 (2019) 109, http://dx.doi.org/10.3390/geosciences9030109.
[64] C. Sammut, G.I. Webb, Encyclopaedia of Machine Learning, Springer, Boston, MA, 2010, http://dx.doi.org/10.1007/978-0-387-30164-8.
[65] B. Herfort, H. Li, S. Fendrich, S. Lautenbach, A. Zipf, Mapping human settlements with higher accuracy and less volunteer efforts by combining crowdsourcing and deep learning, Remote Sensing 11 (2019) 1799, http://dx.doi.org/10.3390/rs11151799.
[66] A. Van Etten, D. Lindenbaum, T.M. Bacastow, SpaceNet: A Remote Sensing Dataset and Challenge Series http://arxiv.org/abs/1807.01232, 2018,. (Accessed 5 May 2020).
[67] D. Chicco, G. Jurman, The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, BMC Genomics 21 (2020) 6, http://dx.doi.org/10.1186/s12864-019-6413-7.
[68] W.F. Vletter, A workflow for (Semi) automatic extraction of roads and paths in forested areas from Airborne Laser Scan data, AARGnews 50 (2015) 33–40.
[69] P. Li, Y. Zang, C. Wang, J. Li, M. Cheng, L. Luo, Y. Yu, Road network extraction via deep learning and line integral convolution, in: International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 1599–1602.
[70] Y. Nachmany, H. Alemohammad, Detecting roads from satellite imagery in the developing world, in: 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 83–89.
[71] B. Pradhan, M. Ibrahim Sameen, An integrated machine learning approach for automatic highway extraction from airborne LiDAR data and orthophotos, in: Laser Scanning Systems in Highway and Safety Assessment: Analysis of Highway Geometry and Safety Using LiDAR, Springer, Cham, 2020, pp. 61–76.
[72] Y. Xu, Z. Xie, Y. Feng, Z. Chen, Road extraction from high-resolution remote sensing imagery using deep learning, Remote Sensing 10 (2018) 1461, http://dx.doi.org/10.3390/rs10091461.
[73] C.E. Shannon, Communication in the presence of noise, Proceedings of the I.R.E. 37 (1949) 10–21.
[74] W. Vletter, The relative chronology of the road network in the Leitha Hills, Siedlungsforschung. Archäologie – Geschichte – Geographie 36 (2019) 367–384.
[75] J.A. Bakker, A note on prehistoric routes on the veluwe and near Uelzen, in: H. Fokkens, B.J. Coles, A.L. van Gijn, J.P. Kleijne, H.H. Ponjee, C.G. Slappendel (Eds.), Between foraging and farming. An extended broad spectrum of papers presented to Leendert Louwe Kooijmans (Analecta Praehistorica Leidensia 40), Leiden University, Leiden, 2008, pp. 281–286.
[76] O. Ronneberger, P. Fischer, T. Brox, in: U-Net: Convolutional Networks for Biomedical Image Segmentation, 2018, https://arxiv.org/abs/1505.04597. (Accessed 15 September 2020).