



# Article Automatic Road Inventory Using a Low-Cost Mobile Mapping System and Based on a Semantic Segmentation Deep Learning Model

Hugo Tardy<sup>1</sup>, Mario Soilán<sup>2,\*</sup>, José Antonio Martín-Jiménez<sup>1</sup> and Diego González-Aguilera<sup>1</sup>

- <sup>1</sup> Department of Cartographic and Terrain Engineering, University of Salamanca, Calle Hornos Caleros 50, 05003 Ávila, Spain
- <sup>2</sup> GeoTECH Group, CINTECX, Universidade de Vigo, Campus Universitario de Vigo, As Lagoas, Marcosende, 36310 Vigo, Spain
- \* Correspondence: msoilan@uvigo.gal

Abstract: Road maintenance is crucial for ensuring safety and government compliance, but manual measurement methods can be time-consuming and hazardous. This work proposes an automated approach for road inventory using a deep learning model and a 3D point cloud acquired by a low-cost mobile mapping system. The road inventory includes the road width, number of lanes, individual lane widths, superelevation, and safety barrier height. The results are compared with a ground truth on a 1.5 km subset of road, showing an overall intersection-over-union score of 84% for point cloud segmentation and centimetric errors for road inventory parameters. The number of lanes is correctly estimated in 81% of cases. This proposed method offers a safer and more automated approach to road inventory tasks and can be extended to more complex objects and rules for road maintenance and digitalization. The proposed approach has the potential to pave the way for building digital models from as-built infrastructure acquired by mobile mapping systems, making the road inventory process more efficient and accurate.

**Keywords:** road inventory; road parameters; mobile mapping system; 3D point cloud processing; deep learning

## 1. Introduction

Roads are the main and most used land communication axes, accounting for 71.8% of inland transport in EU-27 [1]. Road maintenance is therefore of the utmost importance to the relevant management authorities. However, the road infrastructure maintenance investments have been decreasing significantly since 2008, resulting in the deterioration of the road network as well as additional costs, which increase exponentially as road assets deteriorate [1]. To ensure infrastructure conservation and compliance with road users' safety standards [2,3], government and local authorities must regularly measure, map, and inventory road infrastructures. However, manual approaches disrupting traffic are time-consuming and increase the risk of accidents with material or human consequences. Road work zones remain dangerous, which led the European Union Road Federation (ERF) to launch a program in 2014 to raise awareness among governments and standardize road work zones in European countries to increase safety during road works [4].

In response, mobile mapping systems (MMS) have gained popularity, benefiting from being safer and faster than conventional approaches, since they are dynamic remote sensing techniques that do not disturb the traffic flow. In fact, there is a wide range of applications that can use the same acquisition: from road markings detection [5–7], road boundary estimation [8], and automatic retro reflectivity measures on road signs [9,10], to road inventory [11,12]. The prospect of an automatic road inventory workflow paves the way for the digitization of roads, marking a step forward in the possibility of modeling



Citation: Tardy, H.; Soilán, M.; Martín-Jiménez, J.A.; González-Aguilera, D. Automatic Road Inventory Using a Low-Cost Mobile Mapping System and Based on a Semantic Segmentation Deep Learning Model. *Remote Sens.* **2023**, *15*, 1351. https://doi.org/10.3390/ rs15051351

Academic Editors: Massimo Losa and Nicholas Fiorentini

Received: 3 February 2023 Revised: 22 February 2023 Accepted: 24 February 2023 Published: 28 February 2023



**Copyright:** © 2023 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/). existing roads with the emergence of building information modeling (BIM) applications. The standardization of information regarding transportation infrastructure domains is already being studied by buildingSMART to comply with industry foundation classes (IFC) standards [13]. Indeed, roads present a specific challenge as existing road networks exhibit a heterogeneous level of detail and information.

The purpose of this work is to analyze an automated method for road inventory of a 3D point cloud acquired with a low-cost MMS. This method includes the road width, the number and width of individual lanes, the superelevation, the width of the pavement shoulders, and the barrier height when present. Previous works have proposed various methodologies for road inventory: Holgado-Barco et al. proposed a heuristic method based on the segmentation, classification, and extraction of road markings through intensity thresholding and clustering of discontinuous road lines [12]. Then, an estimation of road parameters, except for barrier height, is performed using PCA and specific thresholds. However, it did not use the road centerline and was limited to highway study cases. Vidal et al. [14] performed segmentation and classification of barrier types on 3D point clouds using an intensity threshold to isolate road markings and fit a plane to the road. Then, a DBSCAN clustering allows distinguishing barriers among non-ground points. However, it focuses on security barriers and does not use more complex semantic segmentation applied to different objects. Regarding road tilt, Gargoum et al. performed a detailed estimation of superelevation for water drainage and roadside slopes on road cross sections of an Alaskan highway [15].

These works highlight the importance of the automatic estimation of various elements surrounding road infrastructures. Being able to make an inventory of road assets is a crucial step to building an accurate digital representation of roads in an automated way. In the context of road parametrization, Soilán et al. [16] proposed a method to extract the road centerline and the geometric features of the road in a format compliant with industry foundation classes (IFC) standards. To this end, the 3D point cloud is semantically segmented using a deep learning approach based on Point Transformer [17]. The point cloud is divided into four classes, including road asphalt and markings. Then, the road centerline is extracted, and a robust curvature estimate is applied. The curvature is used to classify the uniformly sampled points into three geometric classes according to the horizon-tal alignment of the road (i.e., straight lines, circulars, and clothoids) whose parameters are calculated.

Che et al. [18] developed a new specific structure called scan pattern grid as a preprocessing step for features extraction and point cloud segmentation. Scan lines are used to reconstruct the vehicle trajectory, then points are projected onto a 2D plan where rows correspond to a scan angle while columns represent a timestamp of the acquisition. This parametrization allows curved roads to be represented as a straight line. However, this method has only been tested on a one-way acquisition and is not directly applicable to a round-trip acquisition.

An essential step in the extraction of information in road applications is to isolate the elements of interest. In this regard, semantic segmentation addresses this problem by assigning a semantic class to each point in the case of a point cloud, allowing operations to be applied only to the desired elements. Semantic segmentation can be achieved by geometric considerations and thresholds, or by machine learning algorithms, such as random sample consensus (RANSAC) and region-growing [19]. Vidal et al. [14] used a scan angle threshold to isolate the road and characterized the density and verticality of the point clouds to extract safety barriers from the road environment.

An increasingly popular approach relies on deep learning models to segment point clouds [20]. Deep learning research on semantic segmentation for point clouds took a large step forward with the release of PointNet in 2016 [21] as the first model able to directly process raw point clouds without the need for additional 2D information or additional transformations. Since then, great efforts have been made to improve results on benchmark datasets, such as S3DIS [17,20–24]. Point Transformer [17] is a modern

architecture published in 2021, based on self-attention layers using a concept analogous to queries, keys, and values to enrich the input with contextual information. This structure has proven effective and increasingly popular in natural processing languages tasks [25] before being successfully applied to point cloud semantic segmentation, achieving state-of-the-art results.

Deep learning has been successfully employed for the segmentation of 3D point clouds of infrastructure features. In [26], road surface objects are segmented to extract features from the road surface and road markings. However, the information is processed using 2D images resulting from the projection of the 3D data. Ma et al. [27] focus on the road pavement, developing a graph convolution network for pavement crack extraction. While the network obtains good results, the applicability is limited to a single road feature. In [28], PointNet++ is used to extract road footprints from airborne LiDAR point clouds in urban areas. The main drawback of airborne data is that it is not possible to extract road parameters or assets that require better data resolution.

In this context, the motivation of this work is to explore the possibilities of the semantic classification of 3D point clouds to support road inventory tasks. To this end, this work proposes the following contributions:

- 1. Exploit a particular version of a deep learning model based on Point Transformer architecture for the semantic segmentation of 3D point clouds of road environments;
- 2. Use the road centerline to divide roads into cross sections and develop an algorithm to build a rectified road model from a round-trip MMS acquisition to facilitate its parametrization;
- 3. Integrate robust methods to estimate road parameters (i.e., road width, lane number and width, road shoulders width and barrier height).

The remaining work is organized as follows. Case study data are presented in Section 2. Section 2 also describes the methodology which consists of three main steps: (1) point cloud segmentation and road centerline, (2) cross sections and construction of the rectified road model, and (3) geometric inventory of the road for each cross section. The results are presented in Section 3 and discussed in Section 4. Finally, the conclusions and future lines of work are presented in Section 5.

#### 2. Materials and Methods

This part is composed of a subsection describing the materials used for acquisition and three methodological subsections, which are schematized in Figure 1. First, a Point Transformer model is trained to perform semantic segmentation on the 3D point cloud. Second, the segmented point cloud is rectified to produce road cross sections that ease the road parameterization process. Finally, different parameters derived from the road layout (lanes, shoulders, barriers, superelevation) are extracted as the output of this methodology.



Figure 1. Workflow of the proposed method.

#### 2.1. Case Study Data

This work uses data acquired with a custom and low-cost MMS in Ávila (Spain) in July 2021 on a 6 km stretch of a conventional road (AV-110, starting in its kilometric point 0) (Figure 2a). The sensor was mounted on a van with 45° tilt, driving at approximately 80 km/h on the closest lane to the road centerline when possible. The laser scanner is a Phoenix Scout Ultra 32 (Figure 2b) equipped with a Velodyne VLP-32C, with 32 laser beams



and horizontal and vertical fields of view of 360° and 40°, respectively. The scan rate of 600,000 measurements per second (PhoenixLidar, 2021) provides dense 3D point clouds.

**Figure 2.** (a) Trajectory of the case study, a 6 km stretch of a road located in Ávila, Spain (red dot in map); (b) Ad-hoc mobile mapping system (MMS).

#### 2.2. Point Cloud Semantic Segmentation and Road Centerline Extraction

This section focuses on the point cloud semantic segmentation using the Point Transformer architecture and the extraction of the road centerline, both of which are used as input for the rest of the work [16].

Semantic segmentation: The road dataset corresponds to a round-trip MMS acquisition to ensure equal density on both sides of the road. Due to the high density in some areas, the point cloud is subsampled with a distance criterion of 3 cm, which results in a dataset of 103 M points. Ground truth data were obtained by manual labeling for the deep learning training. The labeled dataset consists of 3 M points for training and 3.5 M for the test set after subsampling to ensure enough representative data in both of them. The training and test sets are divided into five classes: asphalt, road markings, road signs, barriers, and other, see Figure 3. The class "others" is defined with all the points that do not belong to the other classes to be segmented. Logically, and considering each class defined, we can see that the dataset is unbalanced towards the classes "asphalt" and "other", which each account for 47% of the manually labeled points in the training set, as shown in Table 1.



**Figure 3.** Point cloud labeling on the training set (**a**) and test set (**b**): each point is labeled defining five classes: asphalt (blue), road markings (dark green), road signs (light green), barriers (red), and other (yellow).

	Trai	ining Set	Test Set		
Class	Number of Points	Proportion (in %)	Number of Points	Proportion (in %)	
Asphalt	1,503,027	47	1,609,010	46	
Road markings	110,770	3	154,874	4	
Road signs	16,036	0.5	5498	0.1	
Barriers	71,725	2	8448	0.2	
Other	1,494,874	47	1,719,488	49	
Total	3,196,432	100	3,497,318	100	

Table 1. Class repartition in training and test sets.

Point Transformer is an architecture introduced by Zhao et al. [17]. Supported by this architecture, we designed a deep learning model composed of fives encoders and five decoders, consisting of a variable number of Point Transformer layers followed by a transition down and, respectively, up layer, as represented in Figure 4. Points are grouped through k-nearest neighbors pooling to reduce the cardinality of points by a factor of 4 at each stage of the architecture.



Figure 4. Deep learning model used for the road inventory, being f the laser intensity.

The model is trained for five classes: asphalt, road markings, road signs, barriers, and other. The original weights from the author's training on S3DIS [17,20–24] are used as initial weights to increase generalization capacity. The number of epochs is set to 300, while the learning rate, set to 0.001, is decayed by a factor of 0.1 every 60 epochs with Adam. The batch size is fixed to 32 samples of about 1500 points each.

To account for class imbalance, a weighted cross entropy loss with weights inversely proportional to the number of points in each class in the training set is used. Laser intensity is used as an additional input f for the deep learning model. To augment the data, a random rotation around the Z-axis is applied, as well as a flipping of the positions and slight rotations of up to 15 degrees around the X- and Y-axes for each batch.

Finally, to reduce the classification noise and give a smoother result removing the isolated regions, a conditional random field (CRF) post-processing is added. The energy model defined by Krähenbühl and Koltun [29] is considered as follows (Equations (1) and (2)):

$$E(x) = \sum_{i} \psi_u(x_i) + \sum_{i < j} \psi_p(x_i, x_j)$$
(1)

with

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{i < j} \left( w_1 * \exp\left(-\frac{|p_i - p_j|^2}{2\theta_{\gamma}^2}\right) + w_2 * \exp\left(-\frac{|p_i - p_j|^2}{2\theta_{\alpha}^2} - \frac{|f_i - f_j|^2}{2\theta_{\beta}^2}\right) \right)$$
(2)

where  $p_i$  is the position of the point *i*, its associated class is  $x_i$  and feature  $f_i$ ,  $w_1$  and  $w_2$  are the weights, and  $\theta_{\gamma}$ ,  $\theta_{\alpha}$ ,  $\theta_{\beta}$  are the bandwidths parameters.

The energy Equation (1) is the sum of unaries  $\psi_u$  and pairwise  $\psi_p$  potentials. Unaries potentials correspond to the log probability of the point *i* belonging to a class. The pairwise potentials are composed of a label compatibility matrix  $\mu$  and two kernels, the smoothness kernel and the appearance kernel. Nearby points sharing similar features tend to belong to the same class, which is represented mathematically by the appearance kernel while the smoothness kernel removes small regions in disagreement with their neighbors. The specific features used in the pairwise energy term were selected based on their relevance to this segmentation task, and their effectiveness in reducing classification noise and smoothing the segmentation output was carefully considered, thus, 3D geometric features derived from the coordinates of each point—normal vectors—and intensity as a feature were selected for this term.

Unaries potentials are obtained by inferring the classes scores returned by the deep learning segmentation model. Pairwise potentials allow to consider interactions between points and their associated classes and penalize points classified differently from their neighborhood by defining  $\mu(x_i, x_j) = 1$  if  $x_i \neq x_j$  else 0. By adding a higher energy when points are connected to points of different classes, the minimization of the energy expression *E* results in a smoother segmentation. The minimization is done through a python wrapper of the original code https://github.com/lucasb-eyer/pydensecrf (accessed on 2 February 2023).

However, the hyperparameters of the pairwise energy, the weights  $w_1$ ,  $w_2$ , and the bandwidth  $\theta_{\gamma}$ ,  $\theta_{\alpha}$ ,  $\theta_{\beta}$ , have to be chosen beforehand. To this end, the test set is classified by the final trained model. A grid search of the hyperparameters is performed. For each parameter combination, the increase in mean intersection-over-union (see Section 3) due to the CRF processing is measured on the test set. Experiments showed a small influence of the bandwidth parameters although large values were found to be more beneficial. Optimal values were found for  $w_1 = 0$ ,  $w_2 = 6$  and  $\theta_{\gamma} = 100$ ,  $\theta_{\alpha} = 100$ ,  $\theta_{\beta} = 10,000$ .

Road centerline extraction: The road centerline is extracted following the semantic segmentation of the point cloud. First, points labeled as asphalt and road marking are selected and processed to filter false positives as a post-processing step. Then, road markings are used together with the trajectory obtained by the navigation system of the MMS to extract the road centerline, which separates both traffic directions. That road centerline is processed so it can be defined by continuous geometries associated to the horizontal alignment of the road (i.e., straight lines, circular arcs, and clothoids). Hence, it is possible to sample that road centerline with a fixed distance, obtaining a discrete set of points that belong to the road centerline, which can be used as input data for the next processing steps of this work.

#### 2.3. Cross Sections and Rectified Road Model

This section first describes the methodology for dividing the road point cloud into smaller portions called cross sections, and then the way they are transformed and reattached to build a rectified road model.

Cross-sections: Following a similar approach to Gargoum et al. [15], the road is divided into cross sections. The points of the road centerline are separated by a distance of 1 m and defined by their planimetric coordinates (x, y). A subtraction between consecutive planimetric road centerline points gives the direction vector  $v_i$  of the road on each point  $pos_i$  (Equation (3)).

$$v_i = pos_{i+1} - pos_i = (A \ B \ 0)$$
 (3)

The direction vector  $v_i$ , as a normal vector, and the point  $pos_i = (x, y, 0)$  also define the road cross-section plane Ax + By + Cz + D = 0, visible in Figure 5. The points in the points cloud are then extracted based on a threshold distance *d* to the cross-section plane and a distance to the origin of the vector, which were experimentally chosen as 2.5 m and 20 m, respectively. The distance *d* to the plane is defined as (Equation (4)):

$$d(x, y, z) = \frac{|Ax + By + Cz + D|}{\sqrt{A^2 + B^2 + C^2}}$$
(4)



**Figure 5.** Cross-section plane for the generation of cross sections along the road centerline. This orthogonal plane is defined with the direction vector which performs as the normal vector of the plane and the point from the road centerline.

To facilitate the next steps in the workflow, the road is oriented such that the road axis becomes collinear to the *X*-axis. The angle  $\alpha$  between the direction vector and the *X*-axis is used to define a rotation matrix  $R_{\alpha} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$  which is applied to the cross section in Figure 6.



**Figure 6.** Example of a 10 m cross section: (a) before rotation and (b) after rotation.

Rectified road model: The resulting cross sections are convenient when working individually on each one of them. However, the continuity between consecutive sections is lost. Larger cross sections give more robust lines estimations but are limited by road curves. This step aims at building a rectified model of the road that allows to use the full scale of the road by removing curves. To solve this problem, in addition to the rotation, a translation to align each cross section to the others is performed. The algorithm used in this work requires

only a list of road centerline points and is presented as pseudo-code in Figure 7. A rotation centered on the origin of the associated vector is applied to each section. The section is then translated on the *X*-axis and positioned at the calculated distance between the vector origin and the center of the previous section. The comparison is shown in Figure 8.

Alg	orithm : Cross sections realignement
	Input : trajectory positions pos <sub>i</sub> cross sections associated cloud <sub>i</sub>
	Output : transformed cross sections cloud' <sub>i</sub>
1	Initialization i=1 :
2	$pos'_1 \leftarrow (x_1, 0)$
3	$T_1 = (0, -y_1)$
4	$cloud'_1 = cloud_1 + T_1$
5	Iteration i > 1 :
6	$Distance \leftarrow   pos_i - pos_{i-1}  ^2$
7	$pos'_i \leftarrow (x_{i-1} + Distance, 0)$
8	$T_i \leftarrow pos'_i - pos_i$
9	$cloud'_i \leftarrow cloud_i + T_i$
10	$cloud'_i \leftarrow R_{\theta, pos_i}(cloud'_i)$ with $R_{\theta, pos_i}$ rotation matrix centered on $pos_i$ and $\theta$
	angle between x-axis and pos.

**Figure 7.** Pseudo-code for the algorithm to build a whole rectified model of a curved road from the alignment of each cross section and the road centerline.



**Figure 8.** Comparison of a road subset before and after cross-section alignment. The original point cloud is in the upper part and the rectified road version in the lower part.

## 2.4. Road Parameters

This section focuses on computing the road parameters: road width, superelevation, lanes width, lanes number, road shoulders width, and barriers heights (Figure 9).



Road cross section

Figure 9. Road parameters representation along a road cross section.

First, the cross sections are considered individually to calculate the geometric characteristics shown in Figure 8. Thanks to the rotation computed before (Figure 6), asphalt edges can be approximated as lines whose equation is y = a with *a* constant (Figure 10a). Since the asphalt classification contains false positives and false negatives (see Tables 2 and 3), the outliers are filtered out by approximating the points to a normal distribution along a specific axis and discarding the points outside the range [mean—2\*std, mean + 2\*std]. Filtering is performed on the vertical *Z*-axis and then on the *Y*-axis. Finally, the percentiles 0.01 and 0.99 of the Y distribution of asphalt points are considered as the edges of the asphalt, which are  $y_1$  and  $y_2$ , respectively, with  $y_1 - y_2$  being the road width (Figure 10b). The superelevation can be calculated by using the mean *x* of the cross section and taking the planimetric points closest to  $(x_1, y_1)$  and  $(x_2, y_2)$  belonging to the asphalt class (Equation (5)).

superelevation (%) = 
$$\left(\frac{\Delta Z}{|y_1 - y_2|}\right) x 100$$
 (5)



**Figure 10.** (a) Frequency histogram of points classified as asphalt; (b) rotated cross section with associated 0.01 and 0.99 percentiles which represent the edges of the asphalt.

Metric\Class	Asphalt	Markings	Signs	Other	Barriers	Avg
Precision	0.99	0.74	0.89	0.99	0.76	0.87
Recall	0.96	0.95	0.76	0.99	0.92	0.92
F-score	0.97	0.83	0.82	0.99	0.83	0.89
IoU	0.95	0.71	0.69	0.98	0.72	0.81

Table 2. Class-wise and average metrics of point cloud semantic segmentation.

Table 3. Confusion matrix normalized over the true (rows) classes.

	Asphalt	Markings	Signs	Other	Barriers
Asphalt	0.96	0.03	0	0.01	0
Markings	0.05	0.95	0	0	0
Signs	0	0	0.76	0.12	0.12
Other	0.01	0	0	0.99	0
Barriers	0.02	0.01	0	0.05	0.92

Now that the road width and superelevation are calculated, the lane-related parameters are computed: the number of lanes, the lane width, and the road shoulders width. Their estimation is a process sensitive to outliers and misclassifications, especially when sections are considered individually. To account for this sensitivity, the rectified road model

10 of 20

is used to refine the points classified as road markings. First, road markings points outside the road boundaries (previously computed and represented by  $y_1$  and  $y_2$ ) are discarded.

Even when considered in aggregate, additional road markings such as chevrons, white diagonal stripes (see Figure 11a), and arrows can disrupt the process. Therefore, the entire rectified road marking point cloud is converted to a raster, or more specifically, a binary image. The points are projected onto an XY plane with a resolution of 0.1 m and each pixel containing at least one road marking point has its value set to 1. Then, a Sobel operator [30] defined as follows:

$$S_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$



**Figure 11.** Rasterized binary image of the road markings point cloud: (a) before and (b) after application of Sobel filter to enhance horizontal edges and discarding vertical and diagonal edges and the points outside asphalt edges.

It is applied to all pixels of the image to highlight the horizontal edges. The pixels that do not give a response in 2D (equal to 0) are discarded, while the rest are kept. The result can be seen in Figures 11 and 12. Errors resulting from the alignment process are visible in Figure 12 in the form of undulations and are discussed in more detail in the discussions section.

The next step involves the detection and clustering of road markings using a line extraction based on RANSAC [31]. Haga clic o pulse aquí para escribir texto. RANSAC estimates line parameters by repeatedly using a random sampling strategy. The line parameters are estimated from the subset that contains more inliers, i.e., number of points closer than an orthogonal distance established at 10 cm as threshold (Figure 13).

The lines found on the whole rectified road model, referred to as global lines in the following paragraphs, can then be compared to the lines found with RANSAC in each cross section individually. As central road markings can be separated by a great distance, for each cross section, the two neighboring cross sections are also selected to ensure that at least two central road markings appear. To give a more precise estimation, local lines are estimated using RANSAC and then compared to global lines. Two points are randomly selected from the subset of road markings, a line equation is estimated, and the points



falling within a 0.2 m threshold from the line equation are grouped together, see Figure 14. Only the candidate with the most inliers is retained before repeating the process.

**Figure 12.** Planimetric view of the points classified as road markings along the 1.5 km rectified road model points cloud: (**a**) original and (**b**) after the application of a Sobel filter to discard vertical and diagonal edges and points outside of the road edges.



**Figure 13.** Planimetric view of the points classified as road markings and clustered by the extraction of lines based on RANSAC (left road marking in orange, center road marking in green, right road marking in red, and additional lane in purple).



Figure 14. Planimetric view of cross section road markings with chevron, clustered by colors.

Additional constraints based on the line slope are then added. This allows to discard lines that are diagonal or orthogonal to the road, which can happen in noisy classification or in the presence of chevrons (Figure 14). Each global line found in the rectified road model is associated with the closest local line found in the cross section, which are the ones kept for lanes delimitation. The process is repeated for each cross section.

Although RANSAC can perform well in the presence of outliers, the experimental results applied on individual road sections yielded a high rate of false positives that resulted in erroneous estimates of the width and number of lanes. This is mainly due to the presence of chevrons or white diagonal stripes on the road, in addition to the classification noise. The use of the rectified road model provides a larger scale approach more suitable for lane estimation.

The final step is to calculate the number and the width of lanes. By considering a line defined by the equation x = a, with *a* constant, passing through the center of the road section and the equations of the lines resulting from RANSAC, the number of intersections is calculated. The distances between consecutive lanes can be computed, giving the number and widths of each lane. Similarly, the smallest distances of a road marking to the asphalt edges  $y_1$  and  $y_2$ , correspond to the two road shoulders' widths as represented in Figure 9.

The last part of the method focuses on the calculation of barrier height as seen in Figure 15. Using the classified asphalt points based on the deep learning model defined in Section 2, a plane whose equation is Ax + By + Cz + D = 0 can be fitted to the road with the least squares method. For its part, using the classified barrier points based on the deep learning model defined in Section 2, the barrier height can be computed as the perpendicular distance from the road plane to the barrier. Similar to the road markings, RANSAC allows to extract the line that best fits with the superior part of the barrier and cluster each barrier independently, since the number and height of barriers for each cross section could vary. The 0.95 percentile value along the *Z*-axis is retained as the security barrier height cluster.



**Figure 15.** Cross section computation of barrier height based on the perpendicular distance to the road plane. The two security barriers instances are clustered and treated separately based on RANSAC since their number and height could vary along the cross sections.

#### 3. Results

### 3.1. Point Cloud Segmentation

Following the scientific literature, the mean intersection-over-union (IoU) metric is used to choose the best performing model, which is defined for each class as (Equation (6)). IoU has been chosen instead of other metrics such as overall accuracy, as it is a more informative metric for segmentation tasks when classes are imbalanced, as it is the case for this work:

$$IoU = \frac{TP}{TP + FP + FN} \tag{6}$$

where TP, FP, FN represent the true positives, false positives and false negatives, respectively.

We also compute the following metrics for each class: precision, recall, and f-score. They are resumed in Table 2.

Overall, the model reaches a mean IoU of 0.81. Unsurprisingly, the two most present classes, road and other, have the best semantic segmentation scores. As a consequence, the asphalt class can be considered the most reliable result when used as a basis for further calculations. The confusion matrix in Table 3 shows that most of the confusion for this class also occurs between geometrically similar classes (asphalt/markings). The same confusion occurs between signs and barriers.

CRF post-processing: experiments showed that the performance of the CRF postprocessing increased the mIoU by 0.3. The new values are shown in Table 4 and values in parenthesis indicate the difference without CRF post-processing. Qualitatively, the CRF smooths the classification, hence reducing the classification noise characterized by points isolated from their neighbors when considering their class, as represented in Figure 16. Comparing with similar works in the literature, this work outperforms results from [32] in road surface and guardrail segmentation, or results from [33] in road and traffic sign segmentation.

Metric\Class	Asphalt	Markings	Signs	Other	Barriers	Avg
Precision	0.98 (-0.01)	0.89 (+0.15)	0.94 (+0.05)	0.99 (+0)	0.87 (+0.11)	0.93 (+0.06)
Recall	0.98 (+0.02)	0.90 (-0.05)	0.72 (-0.04)	0.99 (+0)	0.88 (-0.04)	0.89 (-0.03)
F-score	0.98 (+0.01)	0.89 (+0.06)	0.81 (-0.01)	0.99 (+0)	0.88 (+0.05)	0.91 (+0.02)
IoU	0.96 (+0.01)	0.81 (+0.1)	0.69 (+0)	0.98 (+0)	0.78 (+0.06)	0.84 (+0.3)

Table 4. Class-wise and average metrics of point cloud segmentation.

#### 3.2. Road Centerline Extraction

The horizontal road alignment was provided by the local administration as ground truth in a text file, so that it can perform as reference for the extraction of the road centerline. More than 50% of road centerline points present a distance error below 19 cm in comparison

with the ground truth. Larger discrepancies were found to be a result of a discrepancy between the ground truth and the actual road centerline, which could be justified with the complexity and subjective criteria used to define the horizontal alignment of roads or even with discrepancies between the theoretical horizontal alignment designed and the final executed alignment [16].



Figure 16. The classified test set before CRF post-processing (a) and after CRF post-processing (b).

### 3.3. Road Inventory Parameters

The road inventory results are presented in Table 5. To compare the results, a ground truth was collected manually by an expert on a subset of the dataset. Measurements were made manually on the point cloud with the software CloudCompare. The subset is approximately 1.5 km long on which manual measures for different parameters were performed every 20 m.

Point_Id	N° of Lanes	Superelevation (%)	Road Width (m)	Left Shoulder (m)	Right Shoulder (m)	Lane 1 Width (m)	Lane 2 Width (m)	Barrier 1 Height (m)	Barrier 2 Height (m)
285	2	1.6	10.57	1.16	2.10	3.60	3.70		
290	2	0.8	10.58	1.19	2.01	3.57	3.81		
295	2	2.7	10.48	1.18	1.96	3.68	3.66		
300	2	2.0	10.27	1.05	1.91	3.81	3.51	0.14	
305	2	3.4	9.87	1.15	1.53	3.59	3.60	0.12	
310	2	1.6	9.93	1.08	1.63	3.55	3.66		
315	2	4.2	9.81	1.24	1.47	3.55	3.55		
320	2	1.5	9.90	1.24	1.34	3.50	3.81	0.15	
325	2	4.1	10.27	1.53	1.52	3.45	3.77	0.63	0.28
330	2	1.6	10.31	1.58	1.55	3.53	3.65	0.70	0.28
335	2	4.4	10.44	1.61	1.65	3.55	3.63	0.73	
340	2	3.9	10.44	1.40	1.82	3.55	3.67	0.76	
345	2	4.6	10.22	1.47	1.61	3.48	3.66	0.75	0.30

Table 5. Quantitative representation of estimated road parameters using the method developed.

In order to provide a comparison of the road inventory parameters against the ground truth, a graphical representation of the parameters is given in Figure 17. The ground truth is displayed in orange color, whereas the result coming from our method is outlined in blue color. The road width (Figure 17a) and superelevation (Figure 17b) measurements show a good correlation as well as the barriers height (Figure 17f) estimation. However, the correct delimitation of road lanes is a delicate task that results in greater discrepancies in associated parameters (i.e., lanes numbers (Figure 17e), lanes width (Figure 17g), and shoulder widths (Figure 17c,d).



Figure 17. Qualitative analysis of the results for the different road parameters estimated (blue line) and manually measured by an expert (orange line), when available. (a) Road width. (b) Superelevation. (c) Left shoulder width. (d) Right shoulder width. (e) Lanes number. (f) Barrier height. (g) Lanes width.

Quantitative results in Table 6 show the median and mean errors in the estimates of the road inventory parameters. While some road parameters are difficult to estimate and measure manually, such as road shoulders, other parameters such as road width, superelevation, or barrier heights can be estimated with small errors with respect to the ground truth values. Moreover, the number of lanes is correctly inferred in 81% of cases and Figure 17e shows that the parts in which the lane is incorrectly inferred is near singular structures that will be studied further in the discussion section.

Errors	Mean	Median	Mean (%)	Median (%)
Road width (m)/ground truth	0.71	0.35	5	3
Elevation $(\%)/6\%$	0.44	0.36	7	6
Left road shoulder width (m)/1.5 m	0.87	0.24	58	16
Right road shoulder width (m)/1.5 m	0.26	0.23	17	16
Barrier heights (m)/1 m	0.12	0.01	12	1

Table 6. Errors statistics for the road inventory parameters.

Since the measured elements cover different orders of magnitude, a closer look at the errors with respect to the expected value of each element is calculated and represented in the two last columns of Table 6. Values in % represent errors divided by the representative values chosen for each parameter. The road width presents the best result, with a mean error of 0.71 m, representing 5% error. Barrier heights and superelevation are also correctly estimated with a median error percentage of 1% and 6%, respectively, to the expected values. On the contrary, road shoulders present the greatest discrepancy due to their small size and the difficulty to precisely delineate the road edges. It is important to note the high difference between mean and median values in Table 6 due to the presence of outliers and, thus, the robust performance of the median to estimate the errors in the different road inventory parameters. This can be also observed in Figure 17.

## 4. Discussion

In accordance with the results obtained, there are different error sources than can be discussed here. First, the road rectification process takes as reference the extraction of the road centerline. However, this reference centerline contains errors that result in erroneous translations of the cross sections, visible by apparent undulations when displaying the rectified road model. Figure 18a shows the road alignment computed over a subset of points and the resulting road model with a local curb, visible in Figure 18b, that can affect the rest of the process.



(a) Road with centerline error.

(b) Resulting rectified model.

**Figure 18.** Influence of a discrepancy in the road centerline extraction (**a**) on the rectification of the road model (**b**).

Second, the process is highly dependent on the quality of the input semantic segmentation. Large errors propagate into the subsequent heuristic process, resulting in poor parameter estimation. While road asphalt is one of the best segmented classes and is therefore a class that can be considered more reliable than others, this is not the case for road markings whose segmentation strongly influences the traffic lanes delineation or barriers (Figure 19).



**Figure 19.** Semantic segmentation error of a barrier classified as other (yellow) instead of barrier (red) that can hinder the height estimation.

Finally, the road itself may have a variety of singular structures that require special attention to be handled correctly. In the results presented in the previous section, an intersection, visible in Figure 20, consisting of a secondary road joining the main road perpendicularly and a merger lane, have implications for road width, road shoulder width, as well as lanes estimates.



Figure 20. Singular road structures, from left to right: an intersection and a merger lane.

In order to discuss possible improvements of this work, it is interesting to note that in most of these failure cases, there is a geometric logic that is not being fulfilled. For example, in Figure 19, it can be seen how the barrier is not correctly classified on one side of the roadway despite being practically symmetrical with respect to the barrier on the other side, which is correctly classified. Another example is that of erroneous results that may clearly contradict existing regulations regarding shoulder or lane lengths. Adding this type of logic to the classification architecture can improve the results. In this sense, works such as [34] can be of great relevance to add this geometric logic to the segmentation process, limiting these errors thanks to the prior knowledge of the domain.

It is also important to highlight some of the potential applications of the proposed method. First, the straightforward application is road infrastructure mapping: measuring the dimensions and features of the road is critical for road planning and design, as well as for maintenance and safety improvements. The results of this method may be useful to generate standardized as-is information models of the infrastructures, using formats such as IFC [35]. Second, improving the semantic segmentation accuracy and working toward real-time implementations would aid in the development of driver assistance systems, providing crucial information for the perception and decision-making modules of these

systems (for example, lane detection or recognition of traffic signs). Finally, by segmenting the road environment and tracking changes over time, this method can be used to provide information about the deterioration of the road, as well as identify areas where maintenance is required. This can help reduce the cost and time associated with manual inspections and improve the safety and efficiency of the road network, especially if the method is extended to segment geometries related to slopes in mountainous areas, which is a relevant issue for road safety [36].

## 5. Conclusions

This work presents a novel approach for automated road inventory that addresses the challenges of determining road width, number of lanes, lane width, road shoulder width, superelevation, and barrier heights. The approach employs deep learning on 3D point cloud data acquired by a low-cost mobile mapping system (MMS). A deep learning model is designed and trained on a manually labeled subset of the dataset, and the resulting semantic segmentation of the road dataset is refined using a conditional random field (CRF) post-processing to reduce classification noise. Road cross sections are extracted using direction vectors computed from the road centerline, and a rectified road model is generated to aid in lane delineation estimation. The rectified road model is rasterized using a vertical Sobel filter to remove markings diagonal or orthogonal to the road axis, and outliers filtering and heuristic processes are used to estimate the road parameters.

The results of this workflow are compared to a ground truth manually measured by an expert on the point cloud of a 1.5 km-long subset of the road. The estimates yield positive results for road width, superelevation, and barrier heights, with a median error of 0.35 m, 0.36%, and 0.01 m, respectively, and correctly inferring the number of lanes in 81% of the road. This proves the viability of the workflow to support inventory tasks with a more automated and safer approach than the classical protocols used for road inventory.

Although there are potential sources of error that may affect the results, this methodology shows potential for further improvements, such as enhancing the quality of input elements like semantic segmentation, road centerline extraction, and improving the robustness of the heuristic processes to errors. The approach can be refined to extend to more complex objects and rules for road maintenance and digitalization. Adding prior geometric logic to the segmentation network is proposed as an innovative line to improve the presented results. This encourages research on road inventory parameterization, in a context where MMSs and digitalization are increasingly popular. Further research could lead to segmenting more diverse and complex features, paving the way for building digital models from as-built infrastructure acquired by MMS, and to perform more complete geometric assessments.

**Author Contributions:** The road centerline computation was realized by M.S. The semantic segmentation and parameters estimation were realized by H.T. Conceptualization, H.T. and M.S.; Data curation, M.S. and J.A.M.-J.; Formal analysis, H.T.; Investigation, H.T.; Methodology, H.T. and M.S.; Project administration, D.G.-A.; Resources, M.S. and J.A.M.-J.; Supervision, D.G.-A.; Validation, H.T.; Writing—original draft, H.T.; Writing—review and editing, M.S. and D.G.-A. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research leading to these results received funding from the Centro para el Desarrollo Tecnológico Industrial (CDTI) at the INROAD4.0 project (IDI-20181117) under the call 2018 for the strategic program CIEN. This work has been partially supported by Ayuda RYC2021-033560-I611MCIN/AEI of MCIN/AEI/10.13039/501100011033 and European UnionNextGenerationEU/PRTR.

Data Availability Statement: Not applicable.

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

## References

- 1. European Union Road Federation (ERF) an ERF Position Paper for Mantaining and Improving a Sustainable and Efficient Road Network. Available online: http://erf.be/wp-content/uploads/2018/07/Road-Asset-Management-for-web-site.pdf (accessed on 2 February 2023).
- 2. Orden Circular 35/2014, Sobre Criterios de Aplicación de Sistemas de Contención de Vehículos. Available online: http:// normativa.itafec.com/equipamiento-vial/ES.10.05.001.OC.pdf (accessed on 2 February 2023).
- Ministerio de Fomento. Orden FOM/273/2016, de 19 de Febrero, Por La Que Se Aprueba La Norma 3.1-IC Trazado, de La Instrucción de Carreteras; 2016; Volume BOE-A-2016-2217, pp. 17657–17893. Available online: https://www.boe.es/diario\_boe/txt.php?id=BOE-A-2016-2217 (accessed on 2 February 2023).
- 4. Towards Safer Work Zones: A Constructive Vision of the Performance of Safety Equipment for Work Zones Deployed on the TEN-T Roads. Available online: http://www.erf.be/wp-content/uploads/2018/01/Towards\_Safer\_Work\_Zones\_EN\_FINAL.pdf (accessed on 2 February 2023).
- Wang, J.; Zhao, H.; Wang, D.; Chen, Y.; Zhang, Z.; Liu, H. GPS Trajectory-Based Segmentation and Multi-Filter-Based Extraction of Expressway Curbs and Markings from Mobile Laser Scanning Data. *Eur. J. Remote Sens.* 2018, 51, 1022–1035. [CrossRef]
- Rastiveis, H.; Shams, A.; Sarasua, W.A.; Li, J. Automated Extraction of Lane Markings from Mobile LiDAR Point Clouds Based on Fuzzy Inference. *ISPRS J. Photogramm. Remote Sens.* 2020, 160, 149–166. [CrossRef]
- Hata, A.; Wolf, D. Road Marking Detection Using LIDAR Reflective Intensity Data and Its Application to Vehicle Localization. In Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), Qingdao, China, 8–11 October 2014; pp. 584–589. [CrossRef]
- 8. Ma, L. Road Information Extraction from Mobile LiDAR Point Clouds Using Deep Neural Networks. Ph.D. Thesis, University of Waterloo, Waterloo, ON, Canada, 2020.
- 9. Ai, C.; Tsai, Y.J. An Automated Sign Retroreflectivity Condition Evaluation Methodology Using Mobile LIDAR and Computer Vision. *Transp. Res. Part C Emerg. Technol.* **2016**, *63*, 96–113. [CrossRef]
- Soilán, M.; González-Aguilera, D.; del-Campo-Sánchez, A.; Hernández-López, D.; Del Pozo, S. Road Marking Degradation Analysis Using 3D Point Cloud Data Acquired with a Low-Cost Mobile Mapping System. *Autom. Constr.* 2022, 141, 104446. [CrossRef]
- 11. Guan, H.; Li, J.; Cao, S.; Yu, Y. Use of Mobile LiDAR in Road Information Inventory: A Review. *Int. J. Image Data Fusion* **2016**, *7*, 219–242. [CrossRef]
- 12. Holgado-Barco, A.; Riveiro, B.; González-Aguilera, D.; Arias, P. Automatic Inventory of Road Cross-Sections from Mobile Laser Scanning System. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 3–17. [CrossRef]
- 13. BuildingSmart BuildingSmart International. Available online: https://www.buildingsmart.org/ (accessed on 18 May 2020).
- 14. Vidal, M.; Díaz-Vilariño, L.; Arias, P.; Balado, J. Barrier and guardrail extraction and classification from point clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.-ISPRS Arch.* **2020**, *43*, 157–162. [CrossRef]
- 15. Gargoum, S.A.; El-Basyouny, K.; Froese, K.; Gadowski, A. A Fully Automated Approach to Extract and Assess Road Cross Sections From Mobile LiDAR Data. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 3507–3516. [CrossRef]
- 16. Soilán, M.; Tardy, H.; González-Aguilera, D. Deep Learning-Based Road Segmentation of 3D Point Clouds for Assisting Road Alignment Parameterization. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, XLIII-B2-2022, 283–290. [CrossRef]
- Zhao, H.; Jiang, L.; Jia, J.; Torr, P.; Koltun, V. Point Transformer 2021. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 16259–16268.
- Che, E.; Olsen, M.J. An Efficient Framework for Mobile Lidar Trajectory Reconstruction and Mo-Norvana Segmentation. *Remote Sens.* 2019, 11, 836. [CrossRef]
- Vo, A.-V.; Truong-Hong, L.; Laefer, D.F.; Bertolotto, M. Octree-Based Region Growing for Point Cloud Segmentation. *ISPRS J. Photogramm. Remote Sens.* 2015, 104, 88–100. [CrossRef]
- Zhang, J.; Zhao, X.; Chen, Z.; Lu, Z. A Review of Deep Learning-Based Semantic Segmentation for Point Cloud. *IEEE Access* 2019, 7, 179118–179133. [CrossRef]
- Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. Pointnet: Deep Learning on Point Sets for 3d Classification and Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 652–660.
- 22. Landrieu, L.; Simonovsky, M. Large-Scale Point Cloud Semantic Segmentation with Superpoint Graphs. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
- Hu, Q.; Yang, B.; Xie, L.; Rosa, S.; Guo, Y.; Wang, Z.; Trigoni, N.; Markham, A. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11108–11117.
- Thomas, H.; Qi, C.R.; Deschaud, J.-E.; Marcotegui, B.; Goulette, F.; Guibas, L.J. KPConv: Flexible and Deformable Convolution for Point Clouds. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 6411–6420.
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 4 December 2017; pp. 6000–6010.

- 26. Li, H.T.; Todd, Z.; Bielski, N.; Carroll, F. 3D Lidar Point-Cloud Projection Operator and Transfer Machine Learning for Effective Road Surface Features Detection and Segmentation. *Vis. Comput.* **2022**, *38*, 1759–1774. [CrossRef]
- Ma, L.; Li, J. SD-GCN: Saliency-Based Dilated Graph Convolution Network for Pavement Crack Extraction from 3D Point Clouds. Int. J. Appl. Earth Obs. Geoinf. 2022, 111, 102836. [CrossRef]
- Ma, H.; Ma, H.; Zhang, L.; Liu, K.; Luo, W. Extracting Urban Road Footprints from Airborne LiDAR Point Clouds with PointNet++ and Two-Step Post-Processing. *Remote Sens.* 2022, 14, 789. [CrossRef]
- Krähenbühl, P.; Koltun, V. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials 2012. Adv. Neural Inf. Process. Syst. 2011, 24, 109–117.
- 30. Sobel, I.; Feldman, G. An Isotropic 3 × 3 Image Gradient Operator. Sci. Res. 2015, unpublished.
- 31. Fischler, M.A.; Bolles, R.C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]
- Balado, J.; Martínez-Sánchez, J.; Arias, P.; Novo, A. Road Environment Semantic Segmentation with Deep Learning from Mls Point Cloud Data. Sensors 2019, 16, 3466. [CrossRef]
- Vanian, V.; Zamanakos, G.; Pratikakis, I. Improving Performance of Deep Learning Models for 3D Point Cloud Semantic Segmentation via Attention Mechanisms. *Comput. Graph.* 2022, 106, 277–287. [CrossRef]
- 34. Badreddine, S.; d'Avila Garcez, A.; Serafini, L.; Spranger, M. Logic Tensor Networks. Artif. Intell. 2022, 303, 103649. [CrossRef]
- Justo, A.; Soilán, M.; Sánchez-Rodríguez, A.; Riveiro, B. Scan-to-BIM for the Infrastructure Domain: Generation of IFC-Complaint Models of Road Infrastructure Assets and Semantics Using 3D Point Cloud Data. *Autom. Constr.* 2021, 127, 103703. [CrossRef]
- Mateos, R.M.; García-Moreno, I.; Reichenbach, P.; Herrera, G.; Sarro, R.; Rius, J.; Aguiló, R.; Fiorucci, F. Calibration and Validation of Rockfall Modelling at Regional Scale: Application along a Roadway in Mallorca (Spain) and Organization of Its Management. *Landslides* 2016, 13, 751–763. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.