

Article



Lightweight Model Development for Forest Region Unstructured Road Recognition Based on Tightly Coupled Multisource Information

Guannan Lei ¹, Peng Guan ², Yili Zheng ², Jinjie Zhou ¹ and Xingquan Shen ^{3,*}

- ¹ Shanxi Provincial Key Laboratory for Advanced Manufacturing Technology, North University of China, Taiyuan 030051, China; guannanlei@bjfu.edu.cn (G.L.)
- ² School of Engineering, Beijing Forestry University, Beijing 100083, China
- ³ School of Mechanical Engineering, Taiyuan University of Technology, Taiyuan 030051, China
- * Correspondence: 20220100@nuc.edu.cn

Abstract: Promoting the deployment and application of embedded systems in complex forest scenarios is an inevitable developmental trend in advanced intelligent forestry equipment. Unstructured roads, which lack effective artificial traffic signs and reference objects, pose significant challenges for driverless technology in forest scenarios, owing to their high nonlinearity and uncertainty. In this research, an unstructured road parameterization construction method, "DeepLab-Road", based on tight coupling of multisource information is proposed, which aims to provide a new segmented architecture scheme for the embedded deployment of a forestry engineering vehicle driving assistance system. DeepLab-Road utilizes MobileNetV2 as the backbone network that improves the completeness of feature extraction through the inverse residual strategy. Then, it integrates pluggable modules including DenseASPP and strip-pooling mechanisms. They can connect the dilated convolutions in a denser manner to improve feature resolution without significantly increasing the model size. The boundary pixel tensor expansion is then completed through a cascade of two-dimensional Lidar point cloud information. Combined with the coordinate transformation, a quasi-structured road parameterization model in the vehicle coordinate system is established. The strategy is trained on a self-built Unstructured Road Scene Dataset and transplanted into our intelligent experimental platform to verify its effectiveness. Experimental results show that the system can meet real-time data processing requirements (\geq 12 frames/s) under low-speed conditions (\leq 1.5 m/s). For the trackable road centerline, the average matching error between the image and the Lidar was 0.11 m. This study offers valuable technical support for the rejection of satellite signals and autonomous navigation in unstructured environments devoid of high-precision maps, such as forest product transportation, agricultural and forestry management, autonomous inspection and spraying, nursery stock harvesting, skidding, and transportation.

Keywords: DeepLab-Road; expanded convolutional layer; inverted residuals; lightweight model for embedded systems; unstructured road recognition in forest region

1. Introduction

Intelligent driving is dedicated to enhancing the adaptability of unmanned vehicles in intricate and unstructured operating environments, thereby fostering the widespread deployment of autonomous navigation robots in open-world scenarios. As a typical unstructured road in forest areas, the effective identification of forest trails can provide new solutions for the autonomous planning and decision-making of robots in the wild. Consequently, in forest settings with pronounced unstructured features, the inexorable developmental trajectory of autonomous navigation robots entails transcending the limitations imposed by confined environments and extending their applications to a broader spectrum of scenarios [1].



Citation: Lei, G.; Guan, P.; Zheng, Y.; Zhou, J.; Shen, X. Lightweight Model Development for Forest Region Unstructured Road Recognition Based on Tightly Coupled Multisource Information. *Forests* **2024**, *15*, 1559. https://doi.org/10.3390/f15091559

Academic Editors: Tomáš Mikita and Marian Rybanský

Received: 31 July 2024 Revised: 30 August 2024 Accepted: 3 September 2024 Published: 4 September 2024



Copyright: © 2024 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/). The commonly used methods for identifying forest roads are aerial photography and airborne laser scanning (ALS). Aerial photography can be used to locate and measure roads that are clearly distinguishable from the surrounding forest canopy. The use of ALS can facilitate road detection in such situations as well as allow road status and width assessments over large forest areas. However, considering the resolution of aerial images and the huge computational volume of point cloud data, the use of these methods for forest road identification is limited [2]. While forest road identification methods based on convolutional ceural networks (CNNs) have been well developed since 2012, two major challenges remain regarding road image recognition: (i) deep network models focus primarily on improving the accuracy rather than efficiency; (ii) similar to structured driving environments in the United States or Europe, most of the current work is focused on well-structured driving environments, and it cannot be separated from the assistance of high-precision maps [3]. The current solution still has considerable limitations in terms of portability and adaptability in a wider range of forest scenarios and unstructured road conditions.

To address the uncertainty of passable areas for unstructured scenarios, Zhou et al. [4] proposed a dual-space fusion road area extraction method based on color channel enhancement and a gray factor optimization for orchard unstructured road recognition, so as to improve the autonomous navigation ability of the farm picking system. Kausa et al. [5] used single-step and two-step object detector methods to conduct road detection on the unstructured road areas in the open dataset to overcome the influences caused by occlusion, lighting changes, environmental conditions, and viewpoint changes during vehicular operations. Vasavi et al. [6] amalgamated the YOLOv3 and RCNN algorithms for the purpose of vehicle detection and classification in the analysis of high-resolution images, specifically geared towards road target detection during nocturnal conditions. Zhang et al. [7] utilized the multivariate Gaussian and Laplacian of Gaussian (LoG) filters and VGG 16 on high spatial resolution multispectral imagery to extract both primary road and secondary roads in forested areas. However, vanishing point detection is easily misled by cluttered background and is more aptly applied to structured road scenes.

The utilization of Lidar point cloud data is also prevalent in unmanned vehicle systems owing to its inherent benefits in providing intuitive distance measurements and effective obstacle sensing [8,9]. Two-dimensional Lidar is widely used in unmanned vehicles' environment perception due to its simple structure, small number of point clouds, and fast computing speed. Palacín et al. [10] proposed a mobile robot self-localization based on an onboard 2D push-broom Lidar using a reference 2D map previously obtained. Selflocalization with a 2D push-broom Lidar is possible by projecting the scan points in the horizontal plane of the 2D reference map before applying a 2D self-location algorithm. Hassan et al. [11] propose a map correction method with planar environmental constraints, which introduces a complementary filter with moving average filtering for Lidar pose estimation to overcome severe 2D point vibration. Zhang Bin et al. [12] proposed a novel global feature point matching-based loop closure detection algorithm, which created a tightly coupled front-end to mitigate front-end accumulated errors and construct globally consistent maps. In general, high reliability and high confidence perception based on the fusion of 2D Lidar and other sensors remains one of the exploration directions for lightweight models [13].

Direct 3D point cloud semantic segmentation is directly processed by neural network models, such as the MVCNN and SnapNet algorithms [14]. Indirect 3D point cloud semantic segmentation methods include the multi-view-based method and voxel-based method [15,16]. Their essence lies in employing a shared multi-layer perceptron to learn the relevant features of point clouds; solve the problems of non-homogeneity, sparsity, and permutation of point clouds; and improve the efficiency of semantic segmentation of point clouds adeptly capture object shape, size, attitude, and positional information, they exhibit irregularity, disorder, and inconsistent density [17,18]. These methods exhibit relatively high computational complexity [19].

With the swift advancement of multi-sensor fusion technology, the effective integration of 2D images into point clouds remains a difficult problem [20,21]. In view of the development of visual SLAM (simultaneous localization and mapping) algorithms, it is important to integrate the features of Lidar point cloud depth information to improve the stability of attitude estimation and local map construction in unmanned navigation systems [22,23]. However, reconstructing the objective function and calculating residual and iterative optimization based on image features and depth information increase the complexity of the problem [24].

In unstructured forest scenes, intelligent roadside extension structures exhibit imperfections that render them incapable of achieving global positioning. It suffers from an absence of effective reference objects and artificial marks, which are characterized by pronounced nonlinearity and uncertainty. This paper introduces "DeepLab-Road", a lightweight and portable nonstructural road identification strategy. The primary contributions of our study are threefold:

- Aiming at unstructured forest road detection and recognition tasks, a lightweight DeepLab-Road algorithm is proposed. To ensure image segmentation accuracy, the number of parameters in the algorithm was considered, and the portability and realtime operation of the algorithm on a flexible and lightweight forest autonomous navigation platform were fully guaranteed.
- By combining image recognition and 2D Lidar data, a quasi-structured road was
 rapidly constructed based on unstructured road image recognition. Parametric construction of unstructured roads presents a framework for autonomous navigation
 vehicles to develop local map coordinates and conduct independent terrain exploration, particularly in environments where satellite signal rejection is prevalent.
- A high-quality unstructured road dataset, the Unstructured Road Scene Dataset (URSD), compensates for the extreme lack of unstructured road open datasets in the field of autonomous navigation and provides a new communication and learning platform for scholars and artists of autonomous navigation and visual inspection research.

2. Materials and Methods

This section introduces the preliminary experimental preparation and the DeepLab-Road model, detailing the network backbone and optimization strategy. This model strives to balance the accuracy and real-time performance of unstructured semantic road recognition.

2.1. Dataset Construction

In the realm of vision-based autonomous navigation architectures, the creation of extensive and comprehensive datasets encompassing various physical conditions such as lighting, weather, temperature, and environmental noise poses a formidable challenge. Over the past decade, the most illustrious vehicle datasets and benchmarks have emerged, including the BIT Vehicle Dataset [25], Comprehensive Vehicle Dataset [26], KITTI Benchmark Dataset [27], MotorBike7500 [28], CamVid [29], and Cityscape [30], among others. These datasets are meticulously crafted to expedite the development of Intelligent Transportation Systems (ITSs) and enhance the precision of Vehicle Type Classification (VTC) [31,32]. However, their acquisition predominantly focuses on urban structured road scenes with robust infrastructures [33–36], revealing a conspicuous gap in the distribution of sample data for unstructured forest scenes.

Therefore, since the project's inception, we have been consistently engaged in data collection and database construction. The focus of the experimental data collection was primarily on unstructured outdoor scenes. In the dataset construction process, it is crucial to ensure mutual independence between the data and eliminate multiple similar images in the same scene to maximize the inter-class variance of the dataset. We compiled comprises than 8000 images capturing diverse unstructured scenes. These include landscapes such as forests, grasslands, farmlands, parks, and rural roads in different seasons. It establishes

an essential database and platform support for subsequent image processing algorithm development and outdoor experiments using an autonomous navigation platform.

It is worth noting that during the dataset construction process outlined in this article, we received significant inspiration and assistance from a fellow scholar in the construction of the RUGD dataset [37]. Therefore, approximately 783 sample images that met the requirements of the RUGD were selected for URSD construction.

2.2. Experimental Platform Construction

The data collection and algorithm validation are integral components of the experimental platform's construction. To meet the experimental requirements, a self-made Ackermann chassis unmanned ground vehicle (UGV) test platform (simulate the structure of forestry skidders) was designed. Additionally, a test environment meeting the specified conditions was configured. The epigynous machine was implemented on a PC with an Intel Core i7-10875H CPU @ 2.5 GHz and 8 GB of RAM with Linux 16.04 and ROS-kinetic. Two-dimensional Lidar (HOKUYO UST_10LX) was selected to obtain point cloud data. The effective detection distance was 10 m. The scanning angle was 270° . The image collected by the CCD vision camera (LRCP10230_1080P from Zhongweiaoke corporation, Shenzhen, China) was 640 (pixels) × 480 (pixels). Shown in Figure 1 is the self-built UGV for this research. The experimental platform adopted the design structure of wood transport vehicles, with a tractor using the Ackermann chassis structure and the trailer using an unpowered four-wheel trailer. The dimensions of the tractor and trailer were consistent: the length of a single vehicle body was 970 mm; the width was 680 mm.



Figure 1. The self-built field UGV platform equipped with a CCD monocular vision camera and a 2D Lidar. (1) CCD monocular vision camera; (2) 2D Lidar; (3) sensor bracket; (4) gyro; (5) Ackerman chassis; (6) epigynous machine.

In order to ensure the accuracy and efficiency of multi-sensor information fusion, a joint calibration method using a visual camera and 2D Lidar was developed and verified in the experiment (including the calibration of spatial dimension and the calibration of time dimension). Detailed information about the calibration method can be found in our previous research [38,39].

The Lidar and the visual camera were 1.00 (m) and 1.10 (m) away from the ground, respectively. The experiments involved in this research are based on this UGV platform. The external parameters of the 2D Lidar and the camera were calibrated, and the result was as follows:

$$R = \begin{bmatrix} 0.0042155 & -0.84729 & -0.0080126 \\ -0.062474 & 0.0033087 & -0.84730 \\ 0.87894 & 0.0016213 & -0.047311 \end{bmatrix}$$
(1)

$$T = \begin{bmatrix} -0.015349 & 0.031108 & -0.25579 \end{bmatrix}^{T}$$
(2)

2.3. DeepLab-Road Model Architecture

This section introduces the DeepLab-Road model, detailing the network backbone and optimization strategy. This model strives to balance the accuracy and real-time performance of unstructured semantic road recognition.

Unlike structured roads, unstructured road images collected in the forest scenarios exhibit varied textural features, inconspicuous feature clutter, and irregular road shapes and boundaries. It is challenging to extract and identify unstructured road areas because of these nonstructural features. Preliminary experiments revealed that common network models could not be directly applied to unstructured road image segmentation tasks. The existing challenges are primarily reflected in three aspects: (i) traditional networks have complex backbone feature extraction networks and a deeper network structure, leading to potential difficulties in training and deployment; (ii) the long-range contextual information in the image has not been fully utilized, which has been proven highly effective in unstructured road segmentation; (iii) the multi-scale features generated by atrous spatial pyramid pooling in traditional networks have limited feature resolution in the scale axis

In response to these issues, a DeepLab-Road Model was designed in this study. As shown in Figure 2, the model utilizes MobileNetV2 as the backbone network and incorporates stripe pooling and DenseASPP into the encoder. The first layer within these bottleneck blocks employs a pointwise convolution with a convolutional kernel size of 1×1 , thereby enabling dimensionality expansion. The second layer involves a separable deep convolution with a spatial extent of 3×3 . Finally, the concluding layer consists of a convolution with a spatial extent of 1×1 . Additionally, the inverted residual structure can facilitate comprehensive feature extraction and mitigate the risk of gradient loss.

and are not dense enough for the unstructured road segmentation scenario.



Figure 2. DeepLab-Road network framework structure.

Then, a DenseASPP structure was incorporated into the backbone network. It combines void convolution and void space pyramid pooling to create a denser feature pyramid and a larger receptive field. DenseASPP achieves richer feature proportional sampling and information sharing by cascading and densely connecting cavity convolutional layers with different expansion rates. So, it can improve the recognition accuracy of the network, and it is particularly suitable for processing high-resolution images and tasks that require capturing a large range of contextual information.

On this basis, the strip pooling strategy is further introduced to reconsider the formula of spatial pooling, so as to enable the backbone network to model the remote dependency effectively. It collects rich contextual information by utilizing pool operations with different kernel shapes to explore images with complex scenes. For each spatial position in the pooled feature map, it encodes its global horizontal and vertical information, and then uses these encodings to balance its own weights for feature optimization. It can be used as an effective plug-and-play module in existing scene analysis networks.

2.3.1. MobileNetV2 Backbone Network

To enhance the speed of unstructured road recognition and reduce the model size, the lightweight network MobileNetV2 is selected as the backbone for DeepLab-Road. The MobileNetV2 model integrates the benefits of deep separable convolution and the inverted residual structure. Specifically, depth separable convolution offers significant advantages in reducing model parameters. The inverted residual structure is advantageous due to the inclusion of a "skip connection" linking the input of the convolutional layer to the output layer, facilitating comprehensive feature extraction and mitigating the risk of gradient loss. Figure 3 illustrates a MobileNetV2 network tailored for unstructured road detection and recognition tasks in forest scenarios.



Figure 3. MobileNetV2 backbone network.

The MobileNetV2 model predominantly comprises convolutional, bottleneck, and pooling layers. The bottleneck blocks in their cores consist of three integral layers. The first layer within these bottleneck blocks employs a pointwise convolution with a convolutional kernel size of 1×1 , thereby enabling dimensionality expansion. The second layer involves a separable deep convolution with a spatial extent of 3×3 . Finally, the concluding layer consists of a convolution with a spatial extent of 1×1 , which plays a pivotal role in the dimensionality.

The fundamental concept of the algorithm revolves around substituting the complete convolution operator with a decomposed convolution operator and partitioning the singular convolution into two distinct layers. The empirical findings suggest that these layers exhibit comparable performance to traditional convolution, albeit at a discernible cost:

$$C = h_i \times w_i \times d_i \left(k^2 + d_j \right) \tag{3}$$

where $h_i \times w_i \times d_i$ is standard convolution input tensor; $h_j \times w_j \times d_j$ is the output tensor; $k \in R^{k \times k \times d_i \times d_j}$ is the convolutional kernel. Depthwise separable convolution effectively reduces computation compared to traditional layers by a factor of k^2 .

The network architecture comprised an initial convolutional layer with 32 output channels, followed by 19 residual bottleneck layers. The selection of ReLU6 as the nonlinear activation function stems from its robustness in low-precision computations. Except for the first layer, a uniform expansion rate is maintained throughout the network, empirically

falling within the range $r \in [5, 10]$. Given the specific focus on unstructured road area delineation in the context of this study, which excludes the complexities of multi-objective segmentation tasks, a judicious balance between curve performance and computational efficiency guided the selection of an expansion rate r = 5.

In the course of model training, adjustable hyperparameters, such as the input image resolution and width multiplier, play a pivotal role in tailoring the architecture to distinct performance criteria. The dataset maintained an image resolution of 640 (pixels) \times 480 (pixels), with width multipliers spanning from 0.3 to 1.5. This configuration results in network models ranging in size from 5.7 (MB) to 30.9 (MB).

2.3.2. DenseASPP Module Design

In the context of the MobileNetV2 backbone network experiment, the model employed in this study ingested high-resolution images within the autonomous driving domain. To augment the receptive field of the convolutional kernel, an expansion rate is required. In addition, forest unstructured road regions in open scenarios exhibit significant uncertainties and variances, posing a formidable challenge to high-dimensional feature representations where accurate encoding of scale information is crucial. To address the issue of generating features with large receptive fields without compromising the spatial resolution, the DenseASPP structure (Figure 4) was incorporated into DeepLab-Road. This integration enhanced the classification robustness of the model.



Figure 4. DenseASPP description.

DenseASPP can be expressed as follows:

$$y_l = T_{k,q_l}(\{y_{l-1}, y_{l-2}, \cdots, y_0\})$$
(4)

where q_l represents the *l* layer expansion rate and *T* denotes the cascade operation, representing the connection of the output features from the previous layers. As shown in Figure 4, DenseASPP stacks all extended convolutions and creates a dense joint. Information from the previous layer is shared between atrous convolution layers in the form of step connections. Therefore, the neurons that generate intermediate features can encode semantic information at different scales. In this study, DenseASPP used dilation convolutions with dilation rates of 3, 6, 12, and 18.

$$R = (d - 1) \times (k - 1) + k$$
(5)

where *R* represents the actual receptive field and *k* represents the convolutional kernel size. At the same time, in order to control the model size and prevent the convolution layers after fusion from being too wide, 1×1 convolution operation is added before each convolution layer. This allows for the depth of the feature map to be compressed to half the original, and continue to participate in the output size control later on. The parameter size of DenseASPP can be calculated by the following formula:

$$S = \sum_{l=1}^{L} \left[c_l \times \frac{c_0}{2} + \frac{c_0}{2} \times K^2 \times n \right]$$
(6)

where c_0 is the initial input feature number; c_l is the number of input features before the first void convolution; *K* is the size of the convolution kernel; *L* is the number of atrous convolution in DenseASPP and $n = c_0/8$.

In the DenseASPP of the DeepLab-Road model, only convolutions with expansion rates of (3, 6, 12, 18) are used. Previous empirical observations suggest that as the expansion rate of atrous convolutions increases, especially when d > 18, the efficacy of feature extraction experiences a gradual decline. Prudent management of the expansion rate proves advantageous in achieving an equilibrium between receptive field dimensions and feature extraction efficiency. DenseASPP amalgamates the merits of both parallel and serial atrous convolution layers, thereby effectively mitigating the issue of vanishing gradients and constraining parameter proliferation.

2.3.3. Stripe Pooling

In the field of autonomous driving, continuously capturing the strong correlation of contextual information in video images provides a means of improving the generalization of the model. The continuous acquisition of contextual information and its strong correlations within video frames constitute a promising avenue for augmenting model generalization. Given the intricacies of unstructured road recognition tasks, it is imperative to capture long-range contextual information. Stripe pooling has emerged as a crucial strategy for acquiring extensive contextual information in pixelwise prediction tasks, as substantiated by its demonstrated efficacy in empowering backbone networks to effectively model long-range dependencies. Accordingly, the stripe pooling strategy was integrated into the DeepLab-Road (Figure 5).



Figure 5. Schematic illustration of the strip pooling module (Different colors indicate channel brightness values).

The introduced module functions as a plug-and-play component, demonstrating its effectiveness in existing scenario resolution networks. Each position within the output tensor was permitted to establish relationships with multiple positions, both vertically and horizontally, in the input tensor. Through the iterative application of the aggregation process, comprehensive remote dependencies across the entire scenario can be established. Additionally, owing to its fundamental multiplication operation, the proposed spatial position multiplexer (SPM) can be construed as an attention mechanism, facilitating its direct application to any pretrained backbone network without requiring retraining. The expression is as follows:

y

$$y_{i}^{h} = \frac{1}{W} \sum_{0 \le j \le W} x_{i,j} y_{j}^{w} = \frac{1}{H} \sum_{0 \le j \le H} x_{i,j}$$
(7)

$$=\mathbf{y}_{i}^{h}+\mathbf{y}_{j}^{w} \tag{8}$$

$$z = E_w(x, \sigma(f(y))) \tag{9}$$

where $y \in \mathbb{R}^{C \times H}$ is the two-dimensional tensor of the output; $x \in \mathbb{R}^{C \times H \times W}$ is the twodimensional tensor of the input; *H* and *W* are the dimension of stripe pooling windows. $z \in \mathbb{R}^{C \times H \times W}$ is output feature map. *f* represents the convolution with size of 1×1 , σ represents the function of sigmoid, and E_w represents element-wise multiplication.

3. Results

This section introduces the experimental results of image recognition based on DeepLab-Road and the preliminary construction of pseudo-structured roads.

3.1. Experimental Environment Configuration

The unstructured road recognition experiment was conducted on a deep learning server running Ubuntu 16.04 as the operation system. The deep learning servers were equipped with an Intel (R) Core (R) i7-10875H CPU @ 2.5 GHz processor, an NVIDIA GeForce RTX 2060 graphics card with 24 GB memory, 32 GB RAM, and 4 TB hard disk. Pytorch 1.7.1 was used to construct the unstructured road recognition model. Python 3.8.10 was used as the programming language, Pytorch as the framework, CUDA 11.1 as the GPU computing platform, and cuDNN v8.0.4 as the GPU acceleration library. Mini-Batch size was 8. Learning rate was 0.0002. Weight decay Was 0.00001. To ensure algorithmic effectiveness and generalizability, a comprehensive collection of unstructured road data from diverse field scenes was actively pursued.

Figure 6a–d were sourced from the RUGD public dataset, and the remaining figures were derived from our self-constructed URSD dataset. The experimental results show that both subjective and objective considerations indicate significant influence of seasonal variations on the image features within the scene's target regions. Therefore, in the URSD set, the unstructured scene includes three different seasons: spring, summer, autumn, and winter. From another perspective, the URSD dataset also includes pavements with different surface structures: vegetated pavements (Figure 6a), gravel pavements (Figure 6b,c), pavement covered by fallen leaves (Figure 6d,e), dirt roads (Figure 6f–i), flagstone roads (Figure 6j–m), road surface overexposed by camera (Figure 6n,o), some partially melted snow (Figure 6o), and a snow-covered road surface (Figure 6p).



Figure 6. Cont.



Figure 6. Unstructured road scenes collected in the URSD dataset.

3.2. Parameterized Construction of Quasi-Structured Roads

Owing to the complexity of the unstructured scene backgrounds and real-time requirements of the algorithm, the experiment specifically focused on unstructured road areas. It employs a single road label without making specific distinctions or recognizing the semantics of nonroad parts. The road area was considered the foreground, and the nonroad area was considered the background. Therefore, this experiment can be viewed as a single-classification experiment. Figure 6 corresponds directly to Figure 7, where the unstructured road area is highlighted with a red mask.

It is evident from the figure that the unstructured road areas were accurately identified and segmented. Even in relatively challenging environments (Figure 7), the overall segmentation was satisfactory. In Figure 7a–d, there is a large amount of weeds or gravel in the road area, and In Figure 7e, even if approximately two-thirds of the areas on both sides of the road are covered with fallen leaves, the unstructured road areas can still be accurately segmented, and the masked area represents the road area. In Figure 7f–h, there is minimal difference in color space dimensions and visual perception between the masked area and the adjacent region; in Figure 7i,j, the road boundaries are very irregular; in Figure 7k,l, the UGV is placed at the boundary of the road, and the road area is not in the visual center; in Figure 7m, the interplay of sunlight casts shadows from the tree canopy onto the road surface; in Figure 7n,o, obviously, the effectiveness of road boundary planning needs to be improved. Overexposed road surfaces and forest roads covered with snow lost some of their original color and texture features, resulting in a consistent color and texture across the entire ground, which may be fatal to our model.

The pseudo-structured road is generated to solve the problem of an irregular road boundary under image segmentation. The boundary of the mask area is irregular and tortuous, which can cause frequent and drastic changes in the direction control commands during the control process, making it impossible to ensure the stability and comfort of the vehicle tracking system. Such road boundaries directly guiding the navigation of autonomous driving systems may lead to catastrophic consequences. Therefore, it is necessary to perform smooth fitting for pseudo-structured roads. The binary image was derived from the mask image, and the boundary delineation of irregular unstructured roads was expeditiously accomplished using the Sobel edge detection algorithm. Subsequently, the furthest point was employed as the pivotal boundary point for the segregation of the left and right borders. Cubic Bessel curve fitting was performed based on the image coordinates of the sampled points. The green curve represents the fitted road boundary



line, whereas the black curve represents the road centerline calculated based on the road boundary line.

Figure 7. Unstructured road region segmentation and preliminary construction of quasistructured roads.

The DeepLab-Road model demonstrates noteworthy segmentation efficacy in such scenarios. From a quantitative analytical standpoint, the delineation of boundaries for these roads is more explicit compared with their unstructured counterparts, yielding a superior and more precise segmentation outcome. Thus, preliminary construction can be achieved by moving from image segmentation to quasi-structured roads.

It should be noted that the images of forest trails in Figures 6 and 7 correspond one-to-one, meaning that the image in Figure 6 is the original image captured by the vehicle-mounted camera, and Figure 7 shows the corresponding algorithm processing result proposed in this study. The collected images were uniformly 640 (pixels) \times 480 (pixels).

3.3. Reprojection of Images and 2D Point Clouds

The semantic segmentation results of the image lack distance information, so combining point cloud information is necessary to construct complete road information in the vehicle coordinate system. Therefore, quasi-structured road construction is considered to provide local coordinate information for the tracking control of UGV in forest scenarios.

Building upon the joint calibration outcomes of the CCD camera and 2D Lidar, a reprojection of the point cloud onto the image was executed. The distance information of the point cloud is assigned to the corresponding pixel along the fitted boundary line, leading to an extension of the pixel coordinate vectors. Subsequently, the positional information of the road boundary is transformed into the actual local scene coordinate system through a continuous sequence of coordinate transformations involving image, camera, vehicle, and world coordinates.

In Figure 8, the green dashed line represents the road boundary fitted to the image. The set of white points is the horizontal line scanned by Lidar. d_2 is the reference distance for camera and Lidar calibration. d_1 is the width of the road boundary in the image. p_1 is the number of pixels corresponding to d_1 . p_2 is the number of pixels corresponding to d_2 . The mathematical description is as follows:

$$\frac{d_1}{d_2} = \frac{p_1}{p_2} \tag{10}$$

According to the proportional equation, d_2 can be obtained as follows: Despite the utilization of 2D Lidar in the proposed scheme, the dynamic movement of the vehicle ensures that the area covered by the point cloud signal encompasses the entire road surface. The unstructured road boundary was parameterized within the local coordinate system, and a smoothed road centerline was fitted using this parameterization. By leveraging a coordinate conversion algorithm, the parameterized road centerline is directly applied to guide the UGV (details of the quantitative evaluation of quasi-structured roads are presented in the next section).

Figure 8. Schematic diagram of the remapping relationship between 2D Lidar point cloud and segmented images.

4. Discussion

In this section, a quantitative evaluation of unstructured road segmentation and the precision of quasi-structured road fitting is presented to thoroughly assess the effectiveness and feasibility of the research.

4.1. Evaluation and Comparison of Image Segmentation Effect

To comprehensively evaluate the open-scene semantic recognition proposed in this study, six widely used image segmentation algorithms were selected for training and testing on the Unstructured Road Scene Dataset.

As shown in the Figure 9, the first column is the original data sample. The image segmentation methods corresponding to the second to eighth columns are as follows: (a) Original image, (b) SegNet, (c) FCN, (d) U-Net, (e) PSPNet, (f) UperNet, (g) DeepLab V3+, and (h) DeepLab-Road. The last column shows the segmentation result of the proposed algorithm DeepLab-Road. The segmentation outcomes are visually presented in the form





of binary images. The red area is the road area recognized by each algorithm, and the black area is the background area.

Figure 9. Recognition results of unstructured road in RUGD dataset by different algorithms. (**a**) Original image; (**b**) SegNet; (**c**) FCN; (**d**) U-Net; (**e**) PSPNet; (**f**) UperNet; (**g**) DeepLab V3+; (**h**) DeepLab-Road.

The SegNet model was the first model to be compared because it was first used on mobile devices such as FPGA, and scholars have also made lightweight optimizations. It exhibits relatively good performance in structured scenes with clean environments, regular target areas, and consistent background textures. However, in a forest scene and unstructured road environment, the image segmentation effect must be improved.

Compared with the SegNet model, U-Net and DeepLab V3+ significantly improved the segmentation of unstructured road areas. The segmentation results of these two algorithms have fragmented patch areas that are scattered over the entire image. Although this method can improve the accuracy index of image segmentation to a certain extent, it causes disastrous problems in subsequent road boundary extraction, quasi-structured road construction, and autonomous navigation system tracking control. Subjectively, the segmentation results of the other algorithms satisfied the preliminary requirements. Therefore, we will further quantitatively evaluate the image segmentation results according to multiple image segmentation indicators.

The results of the algorithm were compared with those of manual calibration to evaluate the image segmentation effect of the algorithm. Owing to the tedious and timeconsuming task of manually labeling the dataset, 800 images were randomly selected from the test set constructed in this study. These images are manually segmented and calibrated at pixel level using the annotation tool "LableMe" for unstructured road areas. The calibration results were compared with the algorithm fitting results. The comparison results are presented in Table 1.

Model	MPA (%)	MIoU (%)	FWIoU (%)	Parameters (M)	MS (MB)	Ave Time (ms)
SegNet	90.26	85.66	88.25	16.31	124.55	95
FCN	92.13	86.05	88.36	134.27	1095.68	701
U-Net	92.14	87.99	90.12	26.36	201.19	167
PSPNet	92.85	88.74	90.71	51.43	392.76	255
UperNet	93.20	89.08	90.98	126.07	962.62	582
DeepLab V3+	93.49	89.27	91.12	59.34	453.47	247
DeepLab-Road	94.86	89.48	91.18	15.13	58.03	83

Table 1. Quantitative evaluation results of comparative studies on URSD.

The three most representative image accuracy indicators—MPA (mean pixel accuracy), MioU (mean intersection over union), and FWIoU (frequency-weighted intersection over union)—are listed in Table 1. In the statistical sample, the performance of these three indicators for the SegNet model in this dataset was inferior to that of the other algorithms, which is consistent with the intuitive perception shown in Figure 9b. The indices of the other algorithms exceed 91%. Compared with the segmentation results of the DeepLab V3+ algorithm, the algorithm proposed in this study is also slightly superior in the image segmentation task of unstructured road scenes.

To comprehensively assess the real-time performance and portability of the algorithm, we introduce metrics such as the number of model parameters (Parameters) and memory occupancy indexes (MS) in Table 1. Notably, there exist substantial disparities in these two parameters among various models, as delineated in the table. Under comparable image accuracy conditions, the proposed model in this study exhibits a parameter count approximately one-fourth that of DeepLab V3+, with memory consumption amounting to merely 12.79% of DeepLab V3+'s memory footprint.

For embedded system models, a delicate balance between the model segmentation performance and parameter efficiency has critical practical significance. The average processing time per frame (Ave Time in Table 1) for the proposed algorithm was 83 (ms), thereby satisfying the minimum requirement of 12 (frames/s) for continuous video image processing.

4.2. Quasi-Structured Road Parameterization and Error Evaluation

Based on image segmentation, matching between the pixel points and point cloud provides an important guarantee for the parameterization of the road model.

Therefore, it is necessary to further analyze the road detection error under the remapping relationship based on the experimental results. An analysis of the reprojection errors is shown in Table 2.

Distance	E_{up} (pixel)	E_{vp} (pixel)	<i>E_{ud}</i> (m)	<i>E_{vd}</i> (m)	E_{ξ} (m)
1 m–2 m	23.614	3.014	0.075	0.007	0.035
2 m–3 m	27.851	3.537	0.122	0.012	0.050
3 m–4 m	31.926	4.028	0.191	0.022	0.079
4 m–5 m	34.864	4.293	0.242	0.026	0.108

Table 2. Remapping error analysis.

In the table, E_{up} and E_{vp} refer to the *U*-axis and *V*-axis pixel average errors between the road boundary pixel sampling point and the corresponding point cloud, respectively. Because more distant pixels in the image represent greater distances, both the horizontal and vertical mean errors of the road boundary sampling points increased with increasing distance. In contrast, the *U*-axis error of the road boundary was larger. Given the characteristics of the control process, the influence of the lateral error on the vehicle heading angle and stability is particularly evident. E_{ud} and E_{vd} were introduced to provide a more direct reflection of the fitting error and denote the actual errors corresponding to E_{up} and E_{vp} , respectively.

During the control process, the road model and control parameters are dynamically updated. The step size and gait of the first few steps in trajectory planning had the greatest impact on the UGV posture adjustment. In addition, unmanned ground vehicles in complex outdoor environments typically operate at a low speed (v < 1.5 m/s). Therefore, the error analysis range remained within 5 (m) ahead of the UGV. In the experiment, errors were measured at 1 m intervals.

 E_{ξ} refers to the lateral average error of the sampling points along the road centerline. The road centerline is derived through the averaging of sampling points corresponding to the left and right boundaries, akin to achieving mean filtering. This process effectively diminishes the fitting error of the road centerline. The maximum average error of the road centerline within a 5 m range is recorded at 0.108 (m). The average relative error is constrained to not exceed 6%. Most of the proposed road widths are within the range of

2 (m)–4 (m). Consequently, there is a strong alignment between image and radar data, affirming that the road model fulfills the requirements for localization and tracking.

5. Conclusions

Forest roads are important features of the environment that have significant impacts on forestry equipment automation and unmanned forest operation. Aiming at the issue of the insufficient generalization ability of intelligent navigation systems in current forest scenarios, this study proposes a lightweight quasi-structured road recognition and reconstruction scheme, DeepLab-Road, suitable for embedded systems. This algorithm has good real-time performance, strong robustness, and environmental adaptability. It can provide more accurate parameterized road models for unmanned ground vehicles to navigate in unstructured scenes. This technology has important application value in promoting the autonomous navigation of intelligent robots in unstructured scenarios, such as forest scenarios, port and dock cargo handling, urban underground space exploration, harsh battlefield environment investigation, independent picking on farms, and forest nursery care.

This study focuses on autonomous navigation technology for unstructured road scenes in forest scenarios and proposes a lightweight quasi-structured road recognition and reconstruction scheme DeepLab-Road suitable for embedded systems. This model uses MobileNetV2 as the backbone network and integrates DenseASPP and the string pooling plug-and-play module to meet the balance between real-time performance of forest engineering vehicles in outdoor environments and image segmentation accuracy. Combining reprojection technology, the detailed geometric shape and topological information of road boundaries provide accurate and guiding directions for the optimal route for vehicles and humans. At the same time, to some extent, it overcomes the problem of real-time computation difficulties in practical application scenarios caused by the redundancy in 3D point cloud data and the lack of clear and unified point cloud data structure features. The construction of pseudo-structured roads provides a parameterized road model for local UGV navigation lacking satellite signals and high-precision map support.

The main contributions of this study also include the self-built dataset URSD. It mainly includes unstructured road data in open scenarios, and the data consistency of this dataset is good. Eliminating interference from factors such as lane lines, traffic lights, people, and vehicles in traditional unmanned driving datasets as much as possible, it enables scholars and models to focus on unstructured road areas. All data samples come from the collected original images, without involving image rotation, image flipping, noise addition, or other extensions. When new samples appear, the training system can be upgraded and updated by adding them to the training set to expand the samples, train, and update the model. This means that the system has good plasticity, robustness, and portability. The dataset that matches the images and point clouds will be systematically curated, disseminated, and showcased in forthcoming research efforts.

Further improvement and optimization are needed in terms of the model architecture. This model replaces deep neural networks with multiple small parallel networks to achieve similar effects to deep neural networks. Multi-line parallelism requires a sufficient number of independent processing units on the CPU or GPU to meet parallel network requirements. Experiments show that when the number of independent CPU cores or independent GPUs is limited, it significantly limits the computational speed of the model. This indicates that our model has a high dependence on computer hardware performance.

It is noteworthy that the dataset explicitly excludes winter scenes with snow-covered roads. In practice, we did capture some unstructured road scenes post-snowfall during winter. However, the homogeneity of color and texture between the foreground and back-ground in images depicting snow-covered roads poses a challenge for feature extraction. This complexity misguides the trained model and introduces confusion and complications. In subsequent endeavors, we plan to extensively augment the dataset and undertake further research and exploration into this particular issue.

Author Contributions: Conceptualization, G.L. and X.S.; methodology, G.L.; software, P.G.; validation, P.G. and G.L.; formal analysis, G.L.; investigation, J.Z.; resources, Y.Z.; data curation, P.G.; writing—original draft preparation, G.L.; writing—review and editing, X.S.; visualization, Y.Z.; supervision, J.Z.; project administration, J.Z.; funding acquisition, G.L. and X.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Shanxi Provincial Key Laboratory for Advanced Manufacturing Technology Foundation: XJZZ202203; Shanxi Provincial Basic Research Program: 202203021222063; the Key Research and Development Program of Shanxi Province of China: 202202150401018.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The financial support mentioned in the Funding statement is gratefully acknowledged.

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

References

- 1. Guo, J.; Wang, G.; Guan, W.; Chen, Z.; Liu, Z. A feasible region detection method for vehicles in unstructured environments based on PSMNet and improved RANSAC. *Multimed. Tools Appl.* **2023**, *82*, 43967–43989. [CrossRef]
- 2. Morley, I.; Coops, N.; Roussel, J.; Achim, A.; Dech, J.; Meecham, D.; McCartney, G.; Reid, D.; McPherson, S. Updating forest road networks using single photon LiDAR in northern Forest environments. *For. Int. J. For. Res.* **2024**, *97*, 38–47. [CrossRef]
- Li, Q.; Garg, S.; Nie, J.; Li, X.; Liu, R.; Cao, Z.; Hossain, M.S. A highly efficient vehicle taillight detection approach based on deep learning. *IEEE Trans. Intell. Transp. Syst.* 2020, 22, 4716–4726. [CrossRef]
- 4. Zhou, X.; Zou, X.; Tang, W.; Yan, Z.; Meng, H.; Luo, X. Unstructured road extraction and roadside fruit recognition in grape orchards based on a synchronous detection algorithm. *Front. Plant Sci.* **2023**, *14*, 1103276. [CrossRef] [PubMed]
- Kausar, A.; Jamil, A.; Nida, N.; Yousaf, M.H. Two-wheeled vehicle detection using two-step and single-step deep learning models. *Arab. J. Sci. Eng.* 2020, 45, 10755–10773. [CrossRef]
- 6. Vasavi, S.; Priyadarshini, N.K.; Harshavaradhan, K. Invariant feature-based darknet architecture for moving object classification. *IEEE Sens. J.* **2020**, *21*, 11417–11426. [CrossRef]
- 7. Zhang, W.; Hu, B. Forest roads extraction through a convolution neural network aided method. *Int. J. Remote Sens.* **2021**, *42*, 2706–2721. [CrossRef]
- 8. Yu, K.; Xu, C.; Ma, J.; Fang, B.; Ding, J.; Xu, X.; Bao, X.; Qiu, S. Automatic matching of multimodal remote sensing images via learned unstructured road feature. *Remote Sens.* **2022**, 14, 4595. [CrossRef]
- 9. Liu, P.; Zhang, G.; Wang, B.; Xu, H.; Liang, X.; Jiang, Y.; Li, Z. Loss function discovery for object detection via convergence simulation driven search. *arXiv* 2021, arXiv:2102.04700.
- 10. Palacín, J.; Martínez, D.; Rubies, E.; Clotet, E. Mobile Robot Self-Localization with 2D Push-Broom LIDAR in a 2D Map. *Sensors* 2020, 14, 2500. [CrossRef]
- 11. Hassan, M.U.; Das, D.; Miura, J. 3D Mapping for a Large Crane Using Rotating 2D-Lidar and IMU Attached to the Crane Boom. *IEEE Access* 2023, *11*, 21104–21116. [CrossRef]
- 12. Zhang, B.; Peng, Z.; Zeng, B.; Lu, J. 2DLIW-SLAM:2D LiDAR-inertial-wheel odometry with real-time loop closure. *Meas. Sci. Technol.* 2024, *35*, 075205. [CrossRef]
- 13. Miao, R.; Liu, X.; Pang, Y. Design of a mobile 3D imaging system based on 2D LIDAR and calibration with levenberg-marquardt optimization algorithm. *Front. Physic* 2022, *10*, 993297. [CrossRef]
- Wang, Y.; Yang, C.; Hu, M.; Zhang, J.; Li, Q.; Zhai, G.; Zhang, X.P. Identification of deep breath while moving forward based on multiple body regions and graph signal analysis. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; Volume 1, pp. 7958–7962.
- 15. Wang, Y.; Hu, M.; Zhou, Y.; Li, Q.; Yao, N.; Zhai, G.; Zhang, X.; Yang, X. Unobtrusive and automatic classification of multiple people's abnormal respiratory patterns in real time using deep neural network and depth camera. *IEEE Internet Things J.* **2020**, *7*, 8559–8571. [CrossRef]
- 16. Mi, X.; Yang, B.; Dong, Z.; Chen, C.; Gu, J. Automated 3D road boundary extraction and vectorization using MLS point clouds. *IEEE Trans. Intell. Transp. Syst.* 2021, 23, 5287–5297. [CrossRef]
- 17. Erfani, S.; Jafari, A.; Hajiahmad, A. Comparison of two data fusion methods for localization of wheeled mobile robot in farm conditions. *Artif. Intell. Agric.* **2019**, *1*, 48–55. [CrossRef]
- 18. Chen, Y.; Xiong, Y.; Zhang, B.; Zhou, J.; Zhang, Q. 3D point cloud semantic segmentation toward large-scale unstructured agricultural scene classification. *Comput. Electron. Agric.* **2021**, *190*, 106445. [CrossRef]
- Li, S.; Corney, J. Multi-view expressive graph neural networks for 3D CAD model classification. *Comput. Ind.* 2023, 151, 103993. [CrossRef]
- 20. Gao, H.; Zhao, F.; Zhang, Y.; Wan, M. Research on multitask model of object detection and road segmentation in unstructured road scenes. *Meas. Sci. Technol.* 2024, *35*, 065113. [CrossRef]

- 21. Zhang, D.; An, Q.; Feng, X.; Liu, R.; Han, J.; Pan, F. Unstructured road extraction in UAV Images based on lightweight model. *Chin. J. Mech. Eng.* **2024**, *37*, 45. [CrossRef]
- 22. Bai, C.; Zhang, L.; Gao, L.; Peng, L.; Li, P.; Yang, L. Real-time segmentation algorithm of unstructured road scenes based on improved BiSeNet. *J. Real-Time Image Proc.* **2024**, *21*, 91. [CrossRef]
- 23. Buján, S.; Guerra-Hernández, J.; González-Ferreiro, E.; Miranda, D. Forest road detection using Lidar data and hybrid classification. *Remote Sens.* **2021**, *13*, 393. [CrossRef]
- Alam, A.; Singh, L.; Jaffery, Z.A.; Verma, Y.K.; Diwakar, M. Distance-based confidence generation and aggregation of classifier for unstructured road detection. J. King Saud Univ.-Comput. Inf. Sci. 2022, 34, 8727–8738. [CrossRef]
- 25. Dong, Z.; Wu, Y.; Pei, M.; Jia, Y. Vehicle type classification using a semi-supervised convolutional neural network. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2247–2256. [CrossRef]
- Yang, L.; Luo, P.; Loy, C.C.; Tang, X. A large-scale car dataset for fine-grained categorization and verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; Volume 1, pp. 3973–3981.
- Espinosa, J.E.; Velastin, S.A.; Branch, J.W. Motorcycle detection and classification in urban Scenarios using a model based on Faster R-CNN. In Proceedings of the 8th International Conference on Pattern Recognition Systems (ICPRS), Valparaiso, Madrid, Spain, 11–13 July 2017; Volume 1, pp. 1–6.
- 28. Zhang, Y.; Yang, R.; Wang, J.; Chen, N.; Dai, Q. The impact of parameters on semantic segmentation: A case study on the camvid dataset. In Proceedings of the 2021 IEEE 23rd International Conference on High Performance Computing & Communications; 7th International Conference on Data Science & Systems; 19th International Conference on Smart City; 7th International Conference on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), Haikou, China, 20–22 December 2021.
- Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; Volume 1, pp. 1–29.
- Liao, Y.; Xie, J.; Geiger, A. KITTI-360: A novel dataset and benchmarks for urban scene understanding in 2D and 3D. *IEEE Trans.* Pattern Anal. Mach. Intell. 2023, 45, 3292–3310. [CrossRef]
- 31. Alghodhaifi, H.; Lakshmanan, S. Holistic Spatio-Temporal graph attention for trajectory prediction in Vehicle–Pedestrian interactions. *Sensors* 2023, 23, 7361. [CrossRef]
- 32. Baheti, B.; Innani, S.; Gajre, S.; Talbar, S. Semantic scene segmentation in unstructured environment with modified DeepLabV3+. *Pattern Recognit. Lett.* **2020**, *138*, 223–229. [CrossRef]
- 33. Yang, B.; Yang, S.; Wang, P.; Wang, H.; Jiang, J.; Ni, R.; Yang, C. FRPNet: An improved Faster-ResNet with PASPP for real-time semantic segmentation in the unstructured field scene. *Comput. Electron. Agric.* **2024**, 217, 108623. [CrossRef]
- Waga, K.; Tompalski, P.; Coops, N.C.; White, J.C.; Wulder, M.A.; Malinen, J.; Tokola, T. Forest road status assessment using airborne laser scanning. *For. Sci.* 2020, 66, 501–508. [CrossRef]
- 35. Valjarević, A.; Djekić, T.; Stevanović, V.; Ivanović, R.; Jandziković, B. GIS numerical and remote sensing analyses of forest changes in the Toplica region for the period of 1953–2013. *Appl. Geogr.* **2018**, *92*, 131–139. [CrossRef]
- 36. Ahn, J.; Kim, M.; Park, J. Autonomous driving using imitation learning with look ahead point for semi structured environments. *Sci. Rep.* **2022**, *12*, 21285. [CrossRef] [PubMed]
- Wigness, M.; Eum, S.; Rogers, J.G.; Han, D.; Kwon, H. A RUGD Dataset for Autonomous Navigation and Visual Perception in Unstructured Outdoor Environments. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, 3–8 November 2019.
- Yao, R.T.; Zheng, Y.L.; Chen, F.J.; Wu, J.; Wang, H. Research on vision system calibration method of forestry mobile robots. *Int. J. Circuits Syst. Signal Proces.* 2020, 14, 1107–1114. [CrossRef]
- 39. Lei, G.; Yao, R.; Zhao, Y.; Zheng, Y. Detection and modeling of unstructured roads in forest areas based on visual-2D Lidar data fusion. *Forests* **2021**, *12*, 820. [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.