Liwen LIU, Weidong YANG, Ben FEI, 2024. GeeNet: robust and fast point cloud completion for ground elevation estimation towards autonomous vehicles. *Frontiers of Information Technology & Electronic Engineering*, 25(7):938-950.
https://doi.org/10.1631/FITEE.2300388

GeeNet: robust and fast point cloud completion for ground elevation estimation towards autonomous vehicles

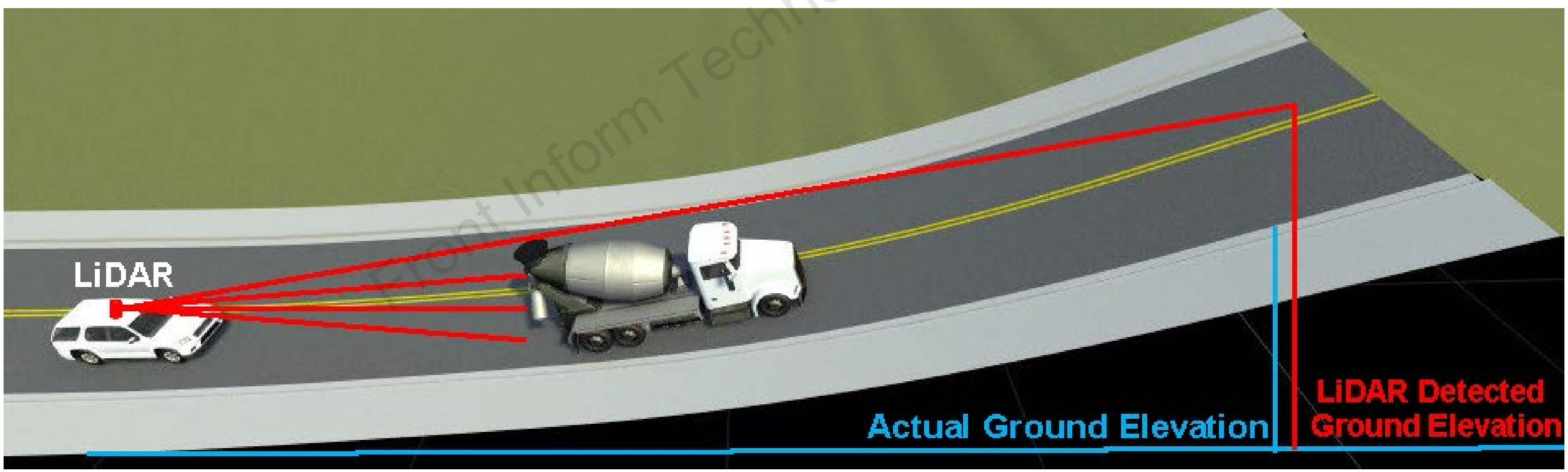
Key words: Point cloud completion; Ground elevation estimation; Real-time; Autonomous vehicles

Corresponding author: Ben FEI
E-mail: bfei21@m.fudan.edu.cn
ORCID: https://orcid.org/0000-0002-3219-9996

Motivation

1. How can we effectively improve complex network architectures and reduce high computing costs?

2. As learning-based approaches require a large amount of annotated data, how can we train the model data without a public dataset for annotated ground elevation maps?



Definition of the actual and detected ground elevation from LiDAR on the car. Due to ground occlusions, LiDAR-detected ground elevation is always different from the actual ground elevation

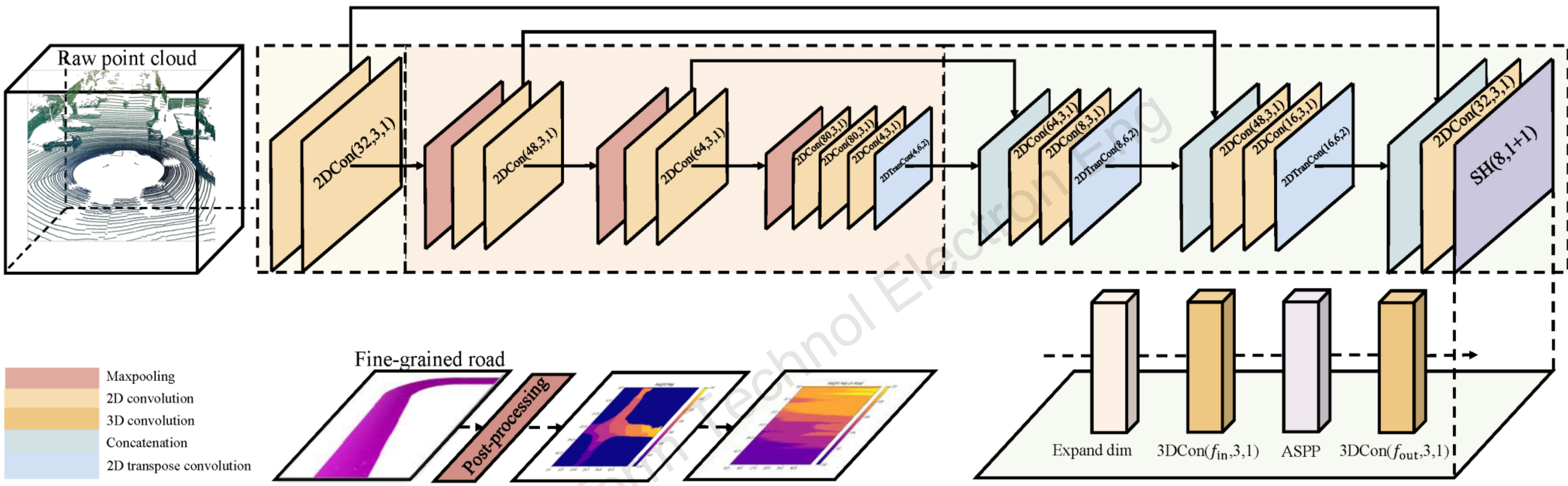
Main idea

1. We propose a novel deep neural network called GeeNet for real-time point cloud completion and ground elevation estimation. GeeNet operates on 3D voxels and is trained in an end-to-end manner. It uses a mixture of 2D/3D convolutions to preserve a lightweight architecture and to regress the ground elevation information of each unit in the grid.

2. The process of generating ground-truth datasets for evaluating ground elevation from the SemanticKITTI and SemanticPOSS datasets is demonstrated experimentally.

3. Comprehensive experiments and comparisons are performed to reveal the robustness, efficiency, and generalization capability of our learning-based GeeNet.

Framework



GeeNet architecture. The pipeline of GeeNet leverages a U-Net along with 2D backbone convolutions (in yellow) and 3D segmentation heads (in purple) to fulfill 3D semantic segmentation and completion with low complexity. Note that convolution parameters are listed as the number of filters, kernel size, and stride. Notice that the dimension of the 2D features is intentionally reduced, and atrous 3D convolutions are integrated to maintain low inference complexity

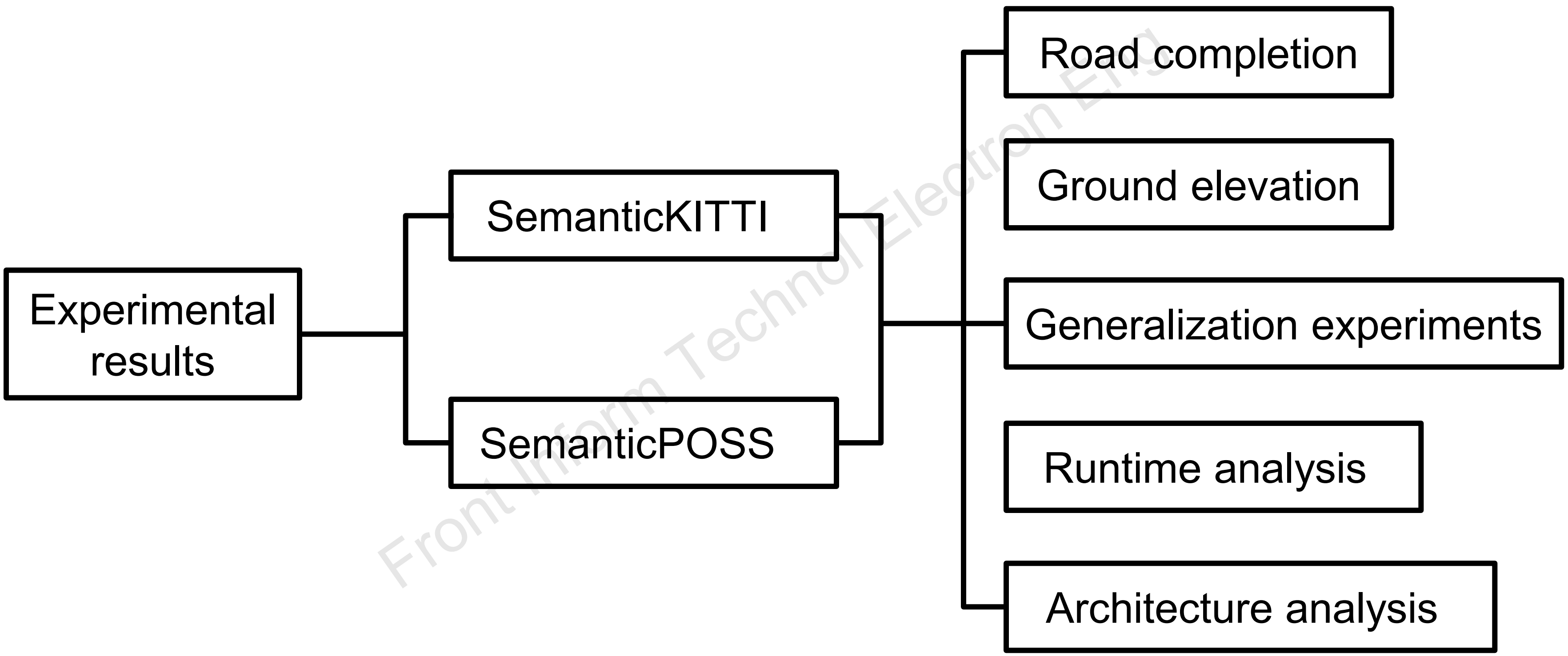
Method

1. We regard the ground elevation estimation problem as a point cloud completion issue, where the height information can be obtained from the complete output of the network by taking a single sparse LiDAR sweep as the input.

2. We cope with the issue of dense 3D semantic completion by assigning a semantic label to each individual voxel.

3. We concatenate the continuous 50 frames, average the z values of all the ground points in the corresponding cells, and create an elevation map. Our goal is to achieve a smooth and uniform ground plane elevation, even in sparse points and occluded areas.

Major results



Major results

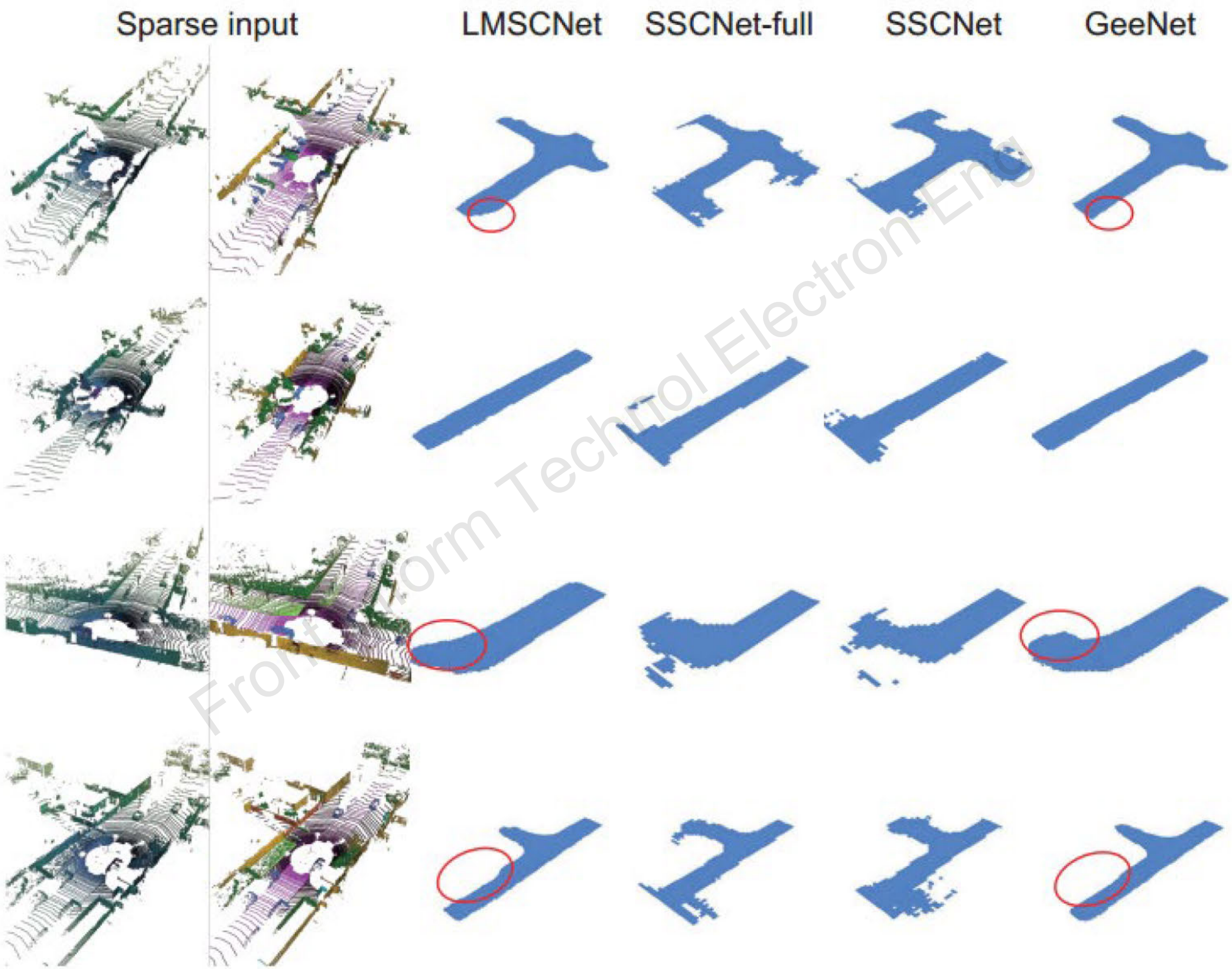


Fig. 4 Comparison of point cloud completion of the road on the SemanticKITTI dataset. The ellipses indicated that GeeNet performed better than LMSCNet, SSCNet-full, and SSCNet

Major results

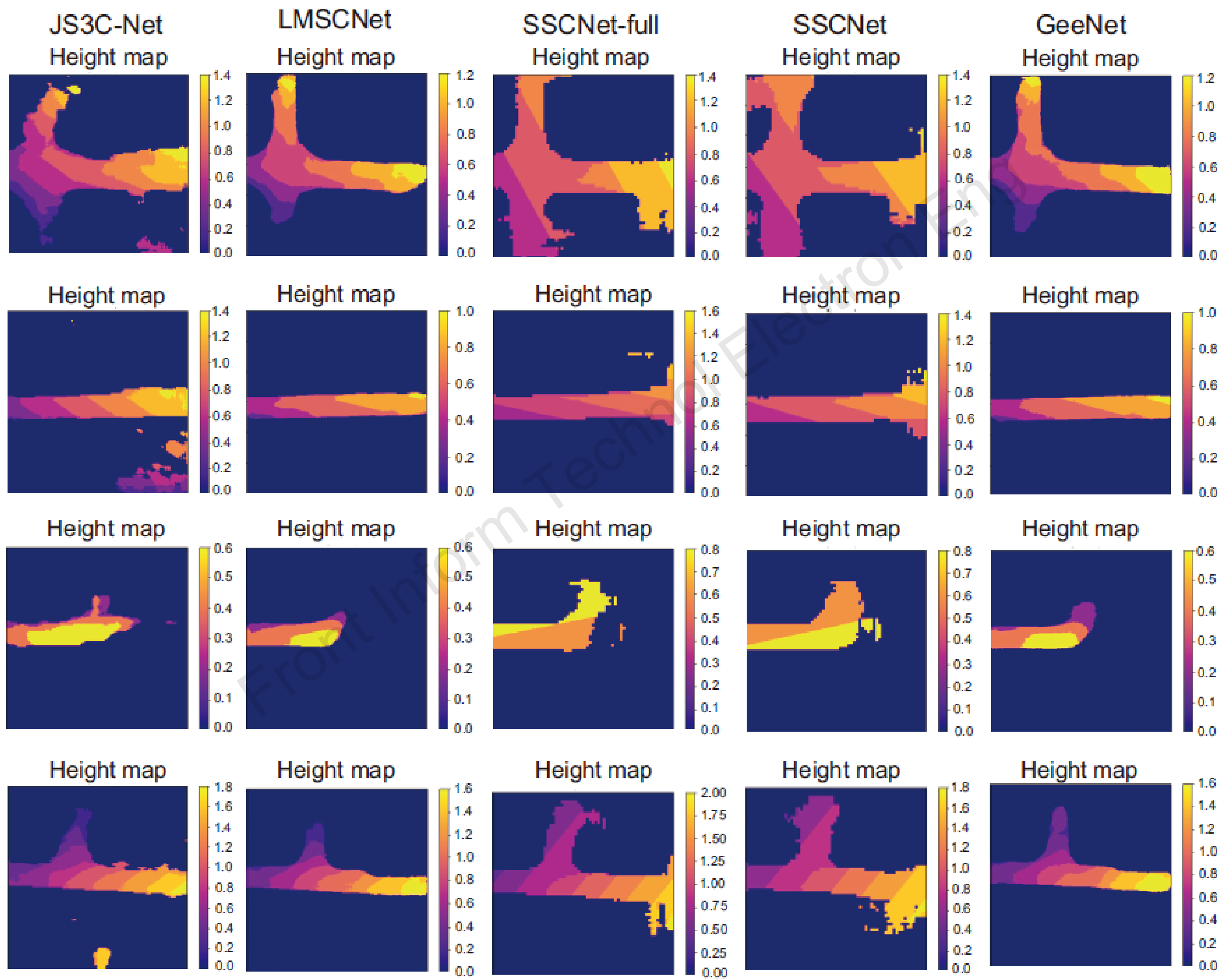


Fig. 5 Comparison of ground elevation estimation on the SemanticKITTI dataset (in meters). GeeNet fulfilled the best ground elevation estimation from the bird's eye view

Major results

Table 2 Quantitative evaluation on point cloud completion of the road on the SemanticKITTI dataset

Method	CD-ℓ_2 ($\times 10^3$)	CD-ℓ_1 ($\times 10^2$)	IoU	Precision	Recall	F-score@1%
JS3C-Net	1.29	0.69	60.75	66.59	77.39	71.36
LMSCNet	1.24	0.62	70.04	83.65	76.95	82.38
SSCNet-full	3.02	3.50	25.45	25.81	64.83	40.57
SSCNet	2.40	1.79	21.62	21.84	65.66	35.56
GeeNet	**0.87**	**0.43**	**71.07**	**85.45**	**80.85**	**83.09**

Table 3 Quantitative evaluation on point cloud completion of the road on the SemanticPOSS dataset

Method	CD-ℓ_2 ($\times 10^3$)	CD-ℓ_1 ($\times 10^2$)	IoU	Precision	Recall	F-score@1%
JS3C-Net	1.65	1.09	55.90	60.11	70.25	67.30
LMSCNet	1.87	1.28	61.34	78.50	72.39	76.21
SSCNet-full	3.89	4.02	22.22	23.19	58.32	36.56
SSCNet	3.35	2.44	20.23	20.39	60.66	33.28
GeeNet	**1.33**	**0.88**	**62.21**	**80.33**	**78.59**	**79.91**

Table 4 Quantitative evaluation of ground elevation estimation on the SemanticKITTI dataset

Method	CD-ℓ_2 ($\times 10^2$)	CD-ℓ_1
JS3C-Net	2.14	0.44
LMSCNet	1.99	0.47
SSCNet-full	2.63	0.54
SSCNet	2.32	0.49
GeeNet	**1.79**	**0.38**

The bold font denotes the best performance

Table 5 Quantitative evaluation of ground elevation estimation on the SemanticPOSS dataset

Method	CD-ℓ_2 ($\times 10^2$)	CD-ℓ_1
JS3C-Net	3.45	1.33
LMSCNet	3.02	1.08
SSCNet-full	3.88	1.82
SSCNet	3.75	1.58
GeeNet	**2.66**	**0.89**

The bold font denotes the best performance

Major results

Table 6 Cross-dataset generalization experiment between SemanticPOSS and SemanticKITTI by GeeNet

Method	IoU	
	SemanticKITTI	SemanticPOSS
GeeNet-KITTI	71.07	58.79
GeeNet-POSS	54.88	62.21

GeeNet trained on SemanticKITTI is denoted as GeeNet-KITTI, while GeeNet trained on SemanticPOSS is denoted as GeeNet-POSS

Table 7 Computational time required by GeeNet and other methods

Method	Time	Device
Rule-based	63.88 s	CPU
JS3C-Net	0.91 ms	GPU*
LMSCNet	0.99 ms	GPU*
SSCNet-full	1.04 ms	GPU*
SSCNet	1.00 ms	GPU*
GeeNet	0.88 ms	GPU*

* RTX 3090

Table 8 Comparison of network statistics among GeeNet and other methods

Method	Number of parameters ($\times 10^6$)	Gflops	FPS
JS3C-Net	2.70	189.8	60.33
LMSCNet	0.35	72.6	21.28
SSCNet-full	1.09	769.6	45.94
SSCNet	0.93	82.5	56.90
GeeNet	**0.33**	**27.3**	**89.30**

The bold font denotes the best performance

Major results

Table 9 Module-level ablation studies on the SemanticKITTI dataset

Model	Upsampling	ASPP	Hierarchical U-Net	IoU
A		√	√	64.90
B	√		√	67.55
C	√	√		68.42
GeeNet	√	√	√	71.07

ASPP: atrous spatial pyramid pooling

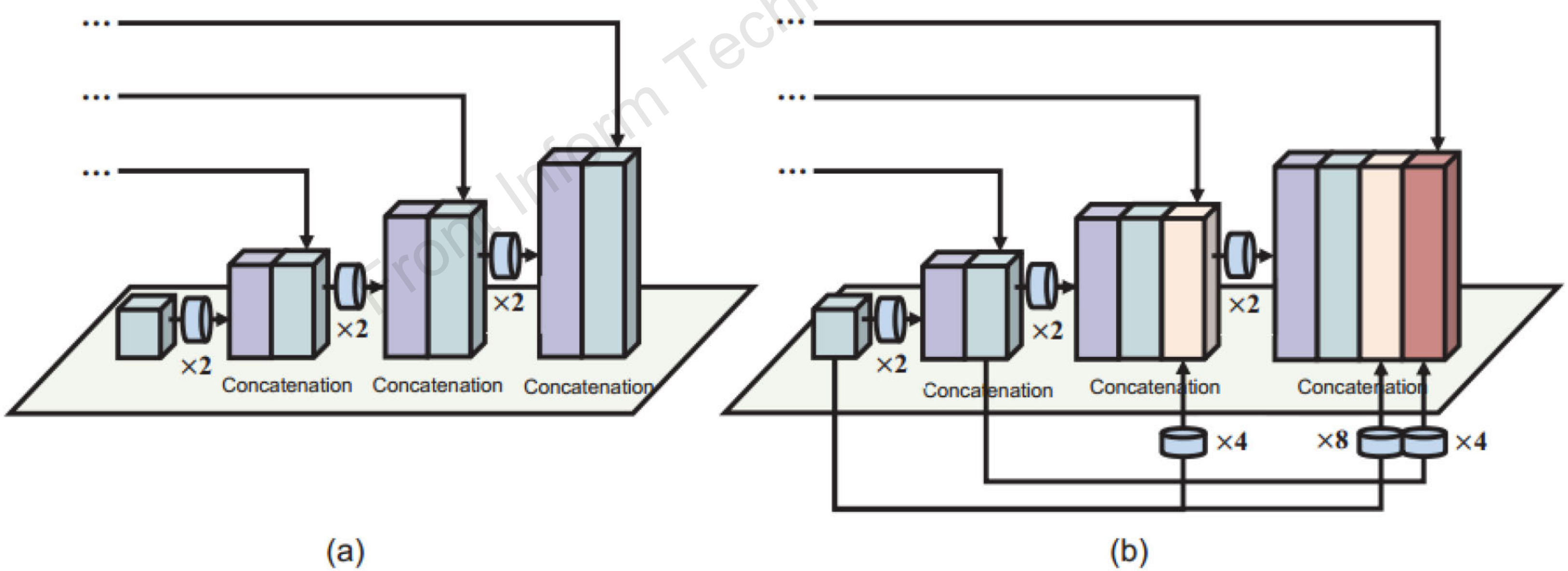


Fig. 6 Architectures of the vanilla U-Net decoder (a) and hierarchical decoder (b)

Conclusions

1. We proposed a robust and fast deep learning based approach called GeeNet, using mixed 2D/3D convolution networks to complete the ground area and output ground elevation in a grid-based representation.

2. The qualitative and quantitative evaluations of GeeNet on the SemanticKITTI and SemanticPOSS datasets demonstrated the superiority of GeeNet to other methods.

3. GeeNet had great generalization ability across different datasets. GeeNet achieved comparable performance in terms of point cloud completion and ground elevation estimation, with a runtime of 0.88 ms.