Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Out-of-distribution detection in 3D semantic segmentation models

Master thesis

September 10, 2021

Lokesh Veeramacheneni

Topics

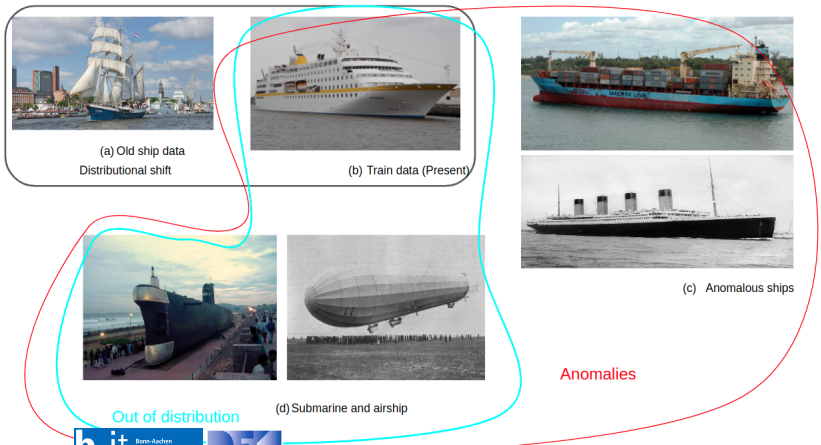
1. Out of distribution (OOD)/Anomaly/Distributional shift
2. Dataset study
3. Model study

OOD/Anomaly/Distributional shift

- **Anomaly** - Patterns that do not conform to the expected behaviour in the training data
- **OOD** - The test input is drawn from an unknown distribution of unknown data which is far from the training distribution
- **Distributional Shift** - The input to the model is a slightly shifted version of training data distribution

OOD/Anomaly/Distributional shift

Figure 1: Images showing the distributional shift in data, anomalies and OOD data using ship example. Images taken from [1], [2], [3], [4], [5] and [6].



Datasets survey

acquisition mode	dataset	frames	points (in million)	classes	scene type
static	Oakland [14]	17	1.6	44	outdoor
	Paris-lille-3D [19]	3	143	50	outdoor
	Paris-rue-Madame [20]	2	20	17	outdoor
	S3DIS [2]	5	215	12	indoor
	ScanObjectNN [23]	-	-	15	indoor
	Semantic3D [10]	30	4009	8	outdoor
	TerraMobilita/IQmulus [24]	10	12	15	outdoor
	TUM City Campus [7]	631	41	8	outdoor
	DALES [25]	40 (tiles)	492	8	outdoor
sequential	A2D2 [8]	41277	1238	38	outdoor
	AIO Drive [26]	100	-	23	outdoor
	KITTI-360 [28]	100K	18000	19	outdoor
	nuScenes-lidarseg [4]	40000	1400	32	outdoor
	PandaSet [27]	16000	1844	37	outdoor
	SemanticKITTI [3]	43552	4549	28	outdoor
	SemanticPOSS [15]	2988	216	14	outdoor
	Sydney Urban [5]	631	-	26	outdoor
	Toronto-3D [21]	4	78.3	8	outdoor
synthetic	SynthCity [9]	75000	367.9	9	outdoor

Figure 2: Surveyed datasets for the 3D LiDAR data.

Semantic3D Vs SemanticKITTI

1. Both are widely evaluated and have highest number of points
2. Both are part of benchmark challenge
3. Synthetic datasets cannot be used for training, they lack the accuracy in detail
4. SemanticKITTI has 3D coordinates and intensity as features
5. In addition to 3D coordinates and intensity, Semantic3D also has RGB
6. Semantic3D is more diverse rather than SemanticKITTI dataset

Training dataset - Semantic3D

1. Large set of points ($\approx$ 4 million)
2. Eight semantic classes
3. Comes with 3D coordinates, RGB and intensity as features
4. Well studied and benchmarked dataset
5. Include variety of urban and rural scenes of various complexity
6. Dense point clouds with little overlap

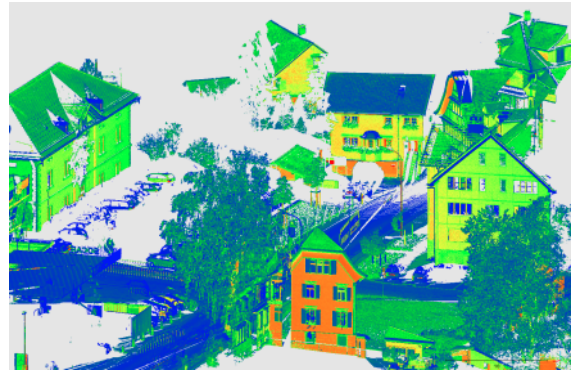


Figure 3: Example of a scene in semantic3D dataset

Models survey

- RandLA-Net has shown higher performance with lower parameters - lower training time

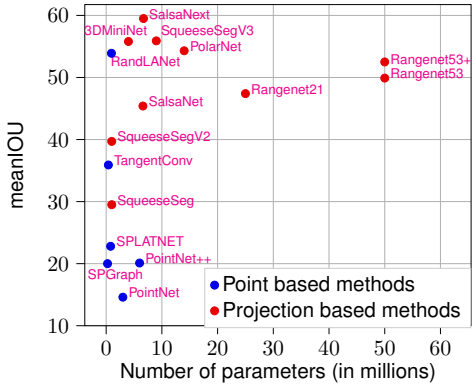


Figure 4: Models performance on SemanticKITTI plotted against the number of parameters

Network - RandLA-Net

1. RandLA-Net employs random point sampling - lower preprocessing time
2. RandLA-Net also uses Dilated residual block - scalable receptive field of a point
3. RandLA-Net also uses attentive pooling - weighted features

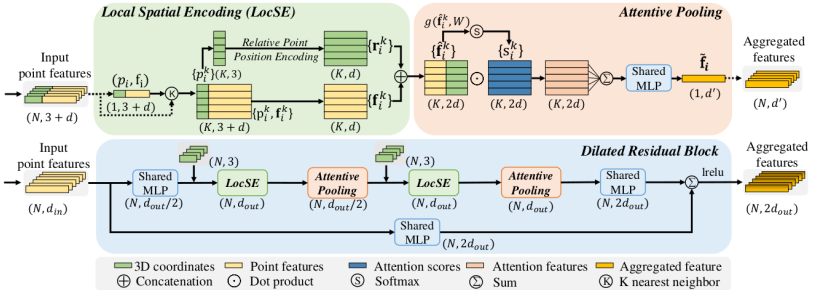


Figure 5: Dilated residual block of RandLA-Net in below image along with subblocks of local spatial encoding and attentive pooling. Image from