

State of the Art This chapter will discuss the 3D LiDAR datasets available and classify them based on the type of . This section will discuss the available LiDAR datasets for the 3D semantic segmentation task and classify the datasets

Static datasets are advantageous because they can produce highly dense point clouds leading to rich geometric representation. The last type of 3D LiDAR dataset is the synthetic dataset. As the name suggests, these datasets are generated from simulation. The datasets belonging to each acquisition type are summed up in Table ?? . Most of the datasets from Table ?? are

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

We chose the Semantic3D dataset as the in-distribution (ID) training dataset from the available datasets. S3DIS is used for 3D semantic segmentation models

This section will discuss the methods available for 3D semantic segmentation. The discussion includes a brief peek into traditional and deep learning approaches. Traditional approach Traditional methods involve a complex features extraction and passing these features to a classifier. Deep learning approach Deep learning based models are efficient at segmentation and can be divided into three types: point-based, projection-based, and graph-based. Let us look briefly into the point-based methods. Point-based methods majorly utilize fully connected layers or tree-based structures. The other flavours of point-based methods include projection onto a d-dimensional lattice or making the point cloud into a graph. In projection-based models, let us discuss the models that project the data onto a 2D range image. The first of this type is the range image projection. We now discuss the projection-based models which involve bird eye view projection. These models are relatively new. Table ?? illustrates the detailed summary of each model and the number of parameters. Finally, graph neural networks are also used for 3D semantic segmentation. Uncertainty estimation methods This section will discuss existing methods to estimate uncertainty in deep neural networks. [?] argues that there are two kinds of uncertainties, they are epistemic and aleatoric uncertainty. Epistemic uncertainty arises due to the model's lack of knowledge about the input data. Because of the higher computational complexity and resource intensiveness, there exist multiple flavours of the deep learning methods for uncertainty estimation. Other methods include snapshot ensembles [?] which iterate over the multiple local optima in the optimization landscape. Bayesian methods Existing neural networks are trained in maximum likelihood, resulting in point estimates for the parameters.

Here θ represents network parameters (weights), $p(\theta)$ represents prior distribution over θ , and x and y represent the input and output respectively.

Here θ represents trained network parameters, x and y represent the training set, and x_t and y_t represent the test set. The Variational Inference (VI) is an approximation method where the posterior probability $p(\theta|x, y)$ is approximated by specific distributions $q(\theta)$. Another widely known example for VI is Monte Carlo Dropout (MC-Dropout), in which the dropout layers are reformed during inference. Out-of-distribution (OOD) detection methods This section will discuss the existing OOD detection methods for 2D image classification. The most widely used benchmark datasets for 2D classification are CIFAR-10 vs SVHN [?], CIFAR-10 vs LSUN [?], and CIFAR-10 vs ImageNet [?].

[?] has proposed a threshold-based OOD detection method using the Mahalanobis distance as a confidence score. The Mahalanobis distance is defined as follows:

where $h(x)$ is final layer activations, c is the constant and $h^-(x)$ is ReAct output of $h(x)$. The score from this ReAct function is used to detect OOD samples. The OOD data samples have higher uncertainty than ID samples, allowing us to use the uncertainty estimation method for OOD detection. In recent years, OOD detection for the task of 2D semantic segmentation has been getting into the limelight. [?] proposed a method for OOD detection in 2D semantic segmentation using the Mahalanobis distance.