2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C)

Survey of State-of-Art Autonomous Driving Technologies with Deep Learning

Yu Huang
Autonomous Driving Research
Black Sesame Technology Inc.
Santa Clara, USA
yu.huang@bst.ai

Yue Chen
Corporate Technology Strategy
Futurewei Technology Inc.
Santa Clara, USA
yue.chen@futurewei.com

***Abstract*— This is a survey of autonomous driving technologies with deep learning methods. We investigate the major fields of self-driving systems, such as perception, mapping and localization, prediction, planning and control, simulation, V2X and safety etc. Due to the limited space, we focus the analysis on several key areas, i.e. 3D object detection, depth estimation from cameras, multiple sensor fusion on the data, feature and task level respectively, behavior modelling and prediction of vehicle driving and pedestrian trajectories.**

Keywords—autonomous driving, deep learning, perception, planning, prediction

I. Introduction

Autonomous Driving has been active for more than 10 years. In 2004 and 2005, DARPA held the Grand Challenges in rural driving of driverless vehicles. In 2007, DAPRA also held the Urban Challenges for autonomous driving in street environments. Then professor S. Thrun at Stanford university, the first-place winner in 2005 and the second-place winner in 2007, joined Google and built Google X and the self-driving team.

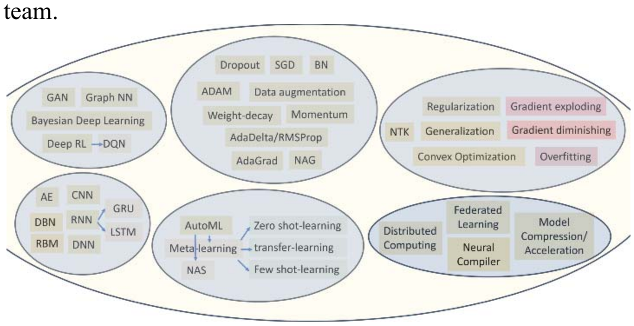

Fig. 1. Deep learning kingdom.

Breakthroughs on deep learning have been achieved since Hinton published new deep structured learning architecture, called deep belief network (DBN) [5]. The past decade has seen rapid developments of deep learning techniques with significant impacts on signal and information processing. In the ImageNet Challenge 2002, the first-place winner came from Hinton's group, using a novel Convolutional Neural Network (CNN) called AlexNet [5].

In this paper, we investigate how autonomous driving marries deep learning [1, 2]. Our survey work spans the state-of-art technology in major fields of self-driving technologies, such as perception, mapping and localization, prediction, planning and control, simulation, V2X and safety etc. Due to the limited space, we focus on some critical areas, i.e. 3D object detection based on different sensors (cameras, radar and LiDAR), depth estimation from cameras, sensor fusion in data, feature and task level respectively, behavior modeling and prediction for vehicle and pedestrian trajectories.

II. Overview of Deep Learning

A. Basic Theory

Like machine learning, deep learning also follows the category as unsupervised, semi-supervised, supervised and reinforcement learning (RL) [5], shown in Fig. 1.

In *supervised learning* domain, there are different deep leaning methods, including Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN) as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). In *unsupervised learning* domain, there are several members for clustering and non-linear dimensionality reduction, including Auto Encoders (AE), Deep Restricted Boltzmann Machines (RBM), and GAN (Generative Adversarial Networks). In addition, RNNs, such as LSTM and Deep RL, are also used for unsupervised learning in many application domains. In *semi-supervised learning* domain, Deep RL and GAN are used; additionally, including RNN (LSTM and GRU) as well.

Deep reinforcement learning is the combination of RL and deep learning [4]. To make machine learning techniques easier to apply and reduce the demand for experienced human experts, *automated machine learning* (AutoML) has emerged as a hot topic [9]. The famous AutoML in deep learning is *neural architecture search* (NAS) proposed by Google [8].

GANs [10] are an unsupervised approach where the generator and the discriminator compete against each other in a zero-sum game. *Graph neural networks* (GNNs) capture the dependence of graphs via message passing between the nodes of graphs [6]. Optimization in training a deep learning model is critical, to avoid overfitting, gradient exploding or diminishing and to accelerate the training process.

There are open deep learning platforms for researchers and engineers to design and develop models, such as PyTorch, Tensorflow, MxNet, Caffe and CNTK.

978-1-7281-8915-4/20/$31.00 ©2020 IEEE
DOI 10.1109/QRS-C51114.2020.00045

B. Distributed Learning

Accelerating deep learning training is a major challenge and techniques range from distributed algorithms to low-level circuit design [7], where a main effort is to exploit their inherent *parallelism*. Most of the operations in learning can be modelled as operations on tensors as a parallel programming model.

C. Model Compression and Acceleration

Deep neural network models are computationally expensive and memory intensive, prohibiting their deployment in devices with small memory resources or in applications with low latency requirements. A solution is to perform model compression and acceleration without significantly decreasing the model performance. So far there are some techniques proposed for use, roughly categorized into four types [3]: parameter pruning and sharing, low-rank factorization, transferred/compact convolutional filters, and knowledge distillation.

Due to the difficulty of deploying various deep learning models on diverse DL hardware, to develop the *deep learning compilers* gets important. Several compilers have been proposed such as Tensorflow XLA/MLIR and the open source TVM [11].

III. Overview of Autonomous Driving

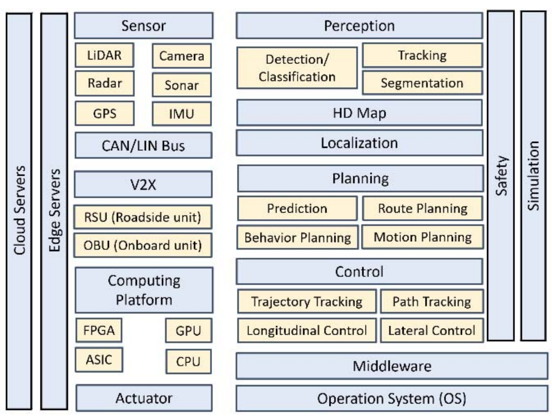

Fig. 2. HW and SW of the autonmous driving platform.

A. Hardware

Fig.2 shows the HW and SW of autonomous driving platform. Autonomous driving vehicle test platforms should be capable of realizing real-time communication, such as in *controller area network* (CAN) buses, and can accurately complete and control the directions, throttles, and brakes of vehicles in real time [15].

Sensing of autonomous driving vehicles falls into three main categories [13]: self-sensing, localization and surrounding sensing. Surrounding sensing uses exteroceptive sensors to perceive road markings, road slope, traffic signs, weather conditions and obstacles. Proprioceptive and exteroceptive sensors can be categorized as either active or passive sensors. The popular sensors include GPS, IMU, cameras, LiDAR, radar and ultrasound etc.

There are different computing platforms, from CPUs, GPUs, ASIC to FPGAs etc., at the vehicle, the roadside and the cloud server.

B. Software

A software platform of autonomous driving is classified multiple layers, from bottom to top as the *real time operating system* (RTOS), middleware, function software and application software. The software architecture could be end-to-end or modular style. Key functions of a modular system are regularly summarized as [13]: perception, localization and mapping, prediction, planning/decision making, and vehicle control etc.

- *Perception* collects information from sensors and discovers relevant knowledge from the environment. It develops a contextual understanding of driving environment, such as detection, tracking and segmentation of obstacles, road signs/marking and free space drivable areas. Based on the sensors implemented, the environment perception task can be tackled by using LIDARs, cameras, radars or a fusion between these three kinds of devices.
- *Mapping* refers to building the map with information of roads, lanes, signs/markings and traffic rules etc. *Localization* determines its position with respect to the driving environment [12].
- *Prediction* refers to estimating the obstacles' trajectories based on their kinematics, behaviors and long-term/short-term histories.
- *Planning* makes decisions on taking the vehicle to the destination while avoiding obstacles, which generates a reference path or trajectory. *Route planning* is referred as finding the point-to-point shortest path in a directed graph. *Behavioral planning* decides on a local driving task that progresses the vehicle towards the destination and abides by traffic rules, traditionally defined by a finite state machine (FSM). *Motion planning* then picks up a continuous path through the environment to accomplish a local driving task, for example RRT and Lattice planning
- *Control* executes the planned actions by selecting appropriate actuator inputs, classified as trajectory or path tracking.
- *V2X (vehicle to everything)* is a vehicular technology system that enables vehicles to communicate with the traffic and the environment around them [14], including vehicle-to-vehicle communication (V2V) and vehicle-to-infrastructure (V2I).

It is worth to mention, the *ISO* (International Organization for Standardization) *26262* standard for *functional safety* of driving vehicles defines a comprehensive set of requirements for assuring safety in vehicle software development [13].

Since driving of an experimental vehicle on the road still costs highly and experiments on existing human driving road networks are restricted, a *simulation* environment is beneficial for developing before real road tests [15].

IV. Perception

In this section, we focus on the detection, reconstruction (depth) and sensor fusion, besides of image processing as denoising and super-resolution, segmentation [110], motion estimation, tracking [111] and human pose estimation (used for pedestrian movement analysis). The detection part is split into 2-D and 3-D. The 3-D method is classified as camera-based, LiDAR-based, radar-based and sensor fusion-based. Similarly, depth estimation is categorized as monocular image-based, stereo and sensor fusion.

The 2-D object detection by deep learning are roughly named as *one-stage* and *two-stage* methods [16]. There are special objects for autonomous driving to detect/classify, i.e. lane [20] and road markings [18], traffic sign [17] and traffic light [19].

A. 3-D Detection

TABLE I. LIDAR-BASED 3-D OBJECT DETECTION METHODS

3-D Volume-based	Projection-based	
	BEV	Frontal view
VoxelNet [22] PointNet[23]/PointNet++ [24] IPOD [30] SECOND [33] PointRCNN [39]/Fast Point RCNN [45] PointPillars [40] Part A2-Net [42] Voxel FPN [43] STD [44] StarNet [46] Sparse 3D CNN [47] VoteNet [48]	RT3D [26] BirdNet [27] HDNet [29] PIXOR [31] YOLO3D [34] Complex-YOLO [35] YOLO4D [36] FaF [38]	Frontal view FCN [21] Grid map CNN [25] LMNet [28] DepthCN [32] Deconvolutional Network [37] FVNet [41] LaserNet [49]

TABLE II. CAMERA-BASED 3-D OBJECT DETECTION METHODS

proposal-based	3D shape-based	2D-3D geometry
Mono3D [50] Deep MANTA [51] Multiple fusion [55] MonoPSR [61] M3D-RPN [66]	Deep MANTA [51] Mono3D++ [57] MonoDIS [62]	Mousavian et al. [52] Joint detection and tracking [53] MonoGRNet [56] GS3D [59] MonoDIS [62] Shift R-CNN [63] Two stage [64] SS3D [65] RTM3D [70] Polygon-Cuboid [71]
Depth map-based	**Transform-based**	**Stereo-based**
Pseudo-LiDAR [58] Pseudo-LiDRA Color [60] ForeSeE [67] RefinedMPL [68]	OFT [54] MoVi-3D [69]	TLNet [73] Stereo R-CNN [72] Pseudo-LiDAR stereo [74] Object centric [75]

For 3-D sensors, like LiDAR and depth sensor (RGB-D), 3-D object detection is direct by finding 3-D bounding box. For single camera, the 3-D object detection needs extensive inference beyond the simple 2-D bound box, to estimate the 3-D bounding box and 3-D pose. Radar can find the object information limited to the scan plane.

1) LiDAR sensors obtain the point cloud data from the surroundings, so the detection methods could be roughly categorized [21-49] as 3-D volume-based and projection-based (mapping 3-D data onto 2-D planes), shown in Table I. Like 2-D detection, the algorithms can fall into one stage and two stage methods too.

2) The camera-based 3-D detection methods can be classified [50-71] as proposal-based, 3D shape-based, 2D-3D geometry-bas ed, depth map-based and 3D transform-based, sjown in Table II.

3) There are some stereo images-based methods with deep learning [72-75].

4) Deep learning has been also applied in radar-based object detection [76-78].

B. Depth Estimation

Depth estimation from images is a reconstruction task in computer vision. Stereo matching [102-106] could be categorized as bottom-up or top-down, 2-D feature-based or 3-D cost volume-based. Depth estimation from monocular image is more challenging than from stereo images. The methods in this domain fall into supervised or unsupervised, with different constraints from edge, surface normal, segment, pose and flow (videos) [79-101], shown in Table III. Depth completion from LiDAR is another topic [107-109].

TABLE III. MONOCULAR CAMERA-BASED DEPTH ESTIMATION METHODS

Unsupervised/semi-supervised/self-supervised Methods			
Stereo	**Camera motion/pose**	**Object motion/flow**	**Normal/Edge**
Garg [83] Godard [84] Kuznietsov [85] Luo [87] Zhou, Fang & Liu [99]	Zhou [86] G Wang [89] Casser [90] Yin & Shi [92] DF-Net [93] Andrahetti [98] Shi [100]	G Wang [89] Yin & Shi [92] DF-Net [93] Casser [96]	LEGO [94]

Supervised Methods				
Brute-force	**Segment/attention**	**Flow**	**Pose/VO**	**Normal**
Eigen, Puhrsch & Fergus [79]	Eigen & Fergus [80] Liu [82] P Wang [83] Jiao [88] CC [95] Bian [97]	CC [95] Casser [96]	C Wang [81] Casser [96]	Eigen&Fergus [80] GeoNet [91]

C. Sensor Fusion

The sensor fusion could be realized in data level and task level. A prerequisite work is calibration of multiple sensors [112-113], to determine transform of aligning the data from different sensors. The sensor fusion methods include depth fusion from camera and LiDAR, and 3-D object detection methods with camera, LiDAR and/or radar.

a) Depth fusion: Similar to depth estimation from images, depth fusion methods with camera images and LiDAR are also categorized [roughly as supervised and unsupervised, with different constraints from pose, flow, edge and surface normal etc [114-131], shown in Table IV.

b) 3-D object detection: Similarly, object detection methods with LiDAR and camera are also classified as volume-based, proposal-based, transform-based and projection-based [132-149], shown in Table V.

V. MAPPING AND LOCALIZATION

In a SLAM survey paper by Cadena et al. [150], *semantic SLAM* is investigated. Semantic mapping consists in associating semantic concepts to geometric entities in robot's surroundings, where deep learning is applied for semantic object detection and classification.

Milz et al. [151] give an overview of deep learning applications in visual SLAM, from depth estimation, optic flow estimation, feature extraction and matching, loop closure detection/re-localization, semantic segmentation and camera pose estimation.

Huang, Zhao and Liu [152] publish an overview of SLAM with LiDAR, camera, IMU and their fusions, in which deep learning methods are investigated in respective sessions too, like feature extraction, object detection, segmentation, moving object removal, pose estimation and localization.

TABLE IV. DEPTH FUSION METHODS WITH LIDAR AND CAMERA

Sparsity-based	Segmentation/Attention	Stereo
Sparse-to-dense [114] Deep Depth Completion [115] Deep Depth Densification [116] Eldesokey, Felsberg and Khan [127]	Jaritz [119] Morphological [120] CFCNet [126]	Park [117] Gansbeke [124] Wang [125]
Image guided	**Motion/Pose**	**Normal**
DPP [122] DFuseNet [123] Tang [128]	DFineNet [129] PLIN [130]	Ma [118] DeepLiDAR [121] Xu [131]

TABLE V. 3-D OBJECT DETECTION METHODS WITH CAMERA AND LIDAR

Volume-based	Projection-based
LiDAR space clustering [133] PointFusion [135] Du [136] Extension of LaserNet [144] Pseudo-LiDAR++ [147]	MV3D [132] Extension of MV3D by tracking [141] Virtual Multi-View Synthesis [145]
Proposal-based	**Transform-based**
AVOD [138] Frustum PointNet [139] SIFRNet [142] CAP [146] MLOD [148] 3D Refinement [149]	Sparse-Nonhomogeneous pooling layer [134] RoarNet [137] Continuous Fusion Layer [140] MVX-Net [143]

VI. PREDICTION, PLANNING AND DECISION MAKING

TABLE VI. PEDESTRAIN BEHAVIOR PREDICTION METHODS

RNN/LSTM	GAN	Attention
Social LSTM [153] Interaction aware motion [154] VRNN [158] SR-LSTM [161] StarNet [162] MATF [163] TraPHic [166] RGM [167] Google Next [168]	Social GAN [155] SoPhie [156] Social Ways [164] IDL [165]	SoPhie [156] Social Attention [157] LVA LSTM[159] ST Attention [160]

TABLE VII. VEHICLE BEHAVIOR MODELING AND DECISION MAKING

CNN-LSTM	GAN/VAE	RL	GNN
Baidu rank-based IRL [169] Berkeley IRL [171] Cui et al. (Deep CNN) [173] TrafficPredict [174] ChauffeurNet [175] CIL [176] INFER [180] Multi-Path [188]	DESIRE [179] AGen (PS-GAIL) [181] Multi-agent [182] CTPS (C-BCGAN) [183] CGNS [184] CVAE + IRL [185] CVAE + STL [186]	Deep RL (DQN) [170] RL-RC [172] MCTC [177] Multi-modal [178]	GRIP [187] Graph conv LSTM [189] Interaction graph [190] Social-WaGDAT [191] EvolveGraph [192] Semantic graph [193] VectorNet [194]

The challenging issues are vehicle/pedestrian behavior modeling and prediction. Ego vehicles will consider learning the driving model, while surrounding vehicles will be predicted for their trajectories [195].

A. *Pedestrian behavior modeling and prediction:* Pedestrian behaviour modelling can be typically classified as physics-based, pattern-based and planning-based. Most of deep learning-based methods and GAN-based methods are pattern-based, while deep reinforcement learning-based methods are planning-based [153-168], shown in Table VI.

B. *Vehcile behavior modeling and decision making:* Vehicle behaviour prediction models are categorized to physics-based, manoeuvre-based, and interaction-aware models. Roughly these methods are calssified based on the model types as CNN, RNN (LSTM, GRU), GAN, GNN and Deep RL/IRL [169-194], shown in Table VII.

VII. END-TO-END SYSTEM

Besides of modular autonomous driving systems, there are some platform working in an end-to-end manner, like a control process [196]. They are either the entire loop from perception to control, loop from planning to control (without perception), or loop from perception to planning (without control) [197-203].

VIII. SIMULATION

Deep learning applications for simulation for autonomous driving mostly fall into sensor modelling, such as radar modelling, LiDAR model and image/video synthesis [204-208].

IX. SAFETY

McAllister et al. [209] investigate three under-explored themes for autonomous driving (AV) research: safety, interpretability, and compliance. A principled approach to modelling uncertainty is Bayesian probability theory, so they propose to use Bayesian deep learning for uncertainty distribution.

VerifAI [210] is a software toolkit from UC Berkeley, for formal design and analysis of systems that include AI and machine learning components. It seeks to address challenges in the presence of environment *uncertainty*. VerifAI provides users with SCENIC, a probabilistic programming language for modelling environments.

X. OPEN DATASETS

There are a number of open data sources (i.e. sensor data, including cameras, LiDAR, radar, GPS/IMU, wheel encoder and so on) in autonomous driving communities, like Kitti [215], Udacity [219], NuScenes [218], Waymo [221], Lyft Level5 [222], BaiduScope [217], BDD (Berkeley) [216], ArgoVerse [220] and PandaSet [224] etc. Some datasets are open for trajectory-based autonomous driving research, like NGSim [225], HighD [226] and INTERACTION [227].

XI. CONCLUSION

We have investigated state-of-art deep learning methods applied for autonomous driving in several major areas. It is seen the marriage of them has made impressive and promising accomplishments. However, there are still some challenges in this field, due to either the autonomous diving task itself or the deep learning shortcomings [211-212], listed as follows.

- *Perception* The "long tail" effect is obvious and there are corner cases to find. To train a deep learning model still requires a lot of data, while model overfitting and sensitivity of image changes are still bother us. In sensor fusion (including V2X as well), modelling of each sensor's capacity and limitation is not well defined.
- *Prediction* The vehicle or pedestrian trajectory prediction needs more data to train the model. More clues are required, extracted from the perception, like human's gaze and pose, drivers' emotion and hand gesture, and vehicles' turn light signal etc.
- *Planning and control* Behavior planning and motion planning are unmatured in deep learning's application, especially real time implementation in crowded and highly dynamic traffic situations. Collaborative planning based on V2X is still a complicated problem.
- *Safety* Uncertainty of deep learning for autonomous driving is still an open problem. Based on GAN's application work [213-214], it is seen some adversarial cases are not easily handled. Fault detection for each module is critical, for fail-safe and fail-operational.
- *Computing platform* It is not clear how the computation power request for autonomous driving is calculated out, especially for planning and control, though there are not a few companies developing stronger SoCs and accelerators.

REFERENCES

[1] S. Grigorescu et al., A Survey of Deep Learning Techniques for Autonomous Driving, arXiv 1910.07738, Mar. 2020

[2] B Ravi Kiran et al., Deep Reinforcement Learning for Autonomous Driving: A Survey, arXiv 2002.00444, Feb. 2020

[3] Y Cheng et al., A Survey of Model Compression and Acceleration for Deep Neural Networks, arXiv 1710.09282, Oct. 2017

[4] Y Li, Deep Reinforcement Learning, arXiv 1810.06339, 2018

[5] M Alom et al., The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches, arXiv 1803.01164, 2018

[6] Z Wu et al., A Comprehensive Survey on Graph Neural Networks, arXiv 1901.00596, Dec. 2019

[7] T Ben-Nun, T Hoefler, Demystifying Parallel and Distributed Deep Learning: An In-Depth Concurrency Analysis, *ACM Computing Surveys*, 9, 2018

[8] Wistuba, M., Rawat, A., Pedapati, T. A Survey on Neural Architecture Search. arXiv 1905.01392, 2019

[9] He X, Zhao K, Chu X, AutoML: A Survey of State-of-Art. arXiv 1908.00709, 2019

[10] Wang Z, Shi Q, Ward T E, Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy, arXiv 1906.01529, 2020.

[11] M Li et al., The Deep Learning Compiler: A Comprehensive Survey, arXiv 2002.03794, April 2020

[12] G Bresson, Z Alsayed, L Yu, S Glaser. Simultaneous Localization and Mapping: A Survey of Current Trends in Autonomous Driving. *IEEE Transactions on Intelligent Vehicles*, 2017

[13] Pendleton S D et al., Perception, Planning, Control, and Coordination for Autonomous Vehicles, *Machines*, 5, 6, 2017

[14] F Jameel et al., Internet of Autonomous Vehicles: Architecture, Features, and Socio-Technological Challenges, arXiv 1906.09918, June 2019

[15] Yurtsever E, Lambert J, Carballo A, Takeda K. A Survey of Autonomous Driving: Common Practices and Emerging Technologies, arXiv 1906.05113, 2020

[16] Zou Z et al., Object Detection in 20 Years: A Survey. arXiv 1905.05055, 2019

[17] Zhu Z et al., Traffic-Sign Detection and Classification in the Wild, IEEE CVPR 2016.

[18] Lee S et al., VPGNet: Vanishing Point Guided Network for Lane and Road Marking Detection and Recognition, arXiv 1710.06288, Oct. 2017

[19] Mueller J and Dietmayer K, Detecting Traffic Lights by Single Shot Detection, arXiv 1805.02523, Oct. 2018

[20] Pan X, Shi J, Luo P, Wang X, Tang X. Spatial as Deep: Spatial CNN for Traffic Scene Understanding, *AAAI*, 2018

[21] Li B, Zhang T, Xia T, Vehicle Detection from 3D Lidar Using FCN, arXiv 1608.07916, 2016

[22] Zhou Y, Tuzel O, VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection, arXiv 1711.06396, 2017

[23] Qi C et al., PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation, arXiv 1612.00593, 2017

[24] Qi C et al., PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space, arXiv 1706.02413, 2017

[25] Wirges S et al, Object Detection and Classification in Occupancy Grid Maps using Deep Convolutional Networks, arXiv 1805.08689, 2018

[26] Zeng Y et al., RT3D: Real-Time 3-D Vehicle Detection in LiDAR Point Cloud for Autonomous Driving, *IEEE RA Letters*, 2018

[27] Beltran J. et al., BirdNet: a 3D Object Detection Framework from LiDAR information, arXiv 1805.01195, 2018

[28] Minemura K et al., LMNet: Real-time Multiclass Object Detection on CPU using 3D LiDAR, arXiv 1805.04902, 2018

[29] Yang B, Liang M, Urtasun R, HDNET: Exploit HD Maps for 3D Object Detection, *CoRL* 2018.

[30] Yang Z et al., IPOD: Intensive Point-based Object Detector for Point Cloud, arXiv 1812.05276, 2018

[31] Yang B, Luo W, Urtasun R, PIXOR: Real-time 3D Object Detection from Point Clouds, arXiv 1902.06326, 2019

[32] Asvadi A et al., DepthCN: Vehicle Detection Using 3D-LIDAR and ConvNet, *IEEE ITSC*, 2017

[33] Yan Y, Mao Y, Li B, SECOND: Sparsely Embedded Convolutional Detection, *Sensors*, 18, 3337, 2018

[34] Ali W et al. YOLO3D: E2E RT 3D Oriented Object Bounding Box Detection from LiDAR Point Cloud, arXiv 1808.02350, 2018

[35] Simon M et al, Complex-YOLO: An Euler-Region-Proposal for Real-time 3D Object Detection on Point Clouds, arXiv 1803.06199, 2018

[36] Sallab A E et al.,YOLO4D: A ST Approach for RT Multi-object Detection and Classification from LiDAR Point Clouds, *NIPS Workshop*, 2018

[37] Vaquero V et al., Deconvolutional Networks for Point-Cloud Vehicle Detection and Tracking in Driving Scenarios, arXiv 1808.07935, 2018

[38] Luo W, Yang B, Urtasun R, Fast and Furious: Real Time E2E 3D Detection, Tracking and Motion Forecasting with a Single Convolutional Net, IEEE CVPR 2018

[39] Shi S, Wang X, Li H, PointRCNN: 3D Object Proposal Generation and Detection from Point Cloud, arXiv 1812.04244, 2018

[40] Lang A et al., PointPillars: Fast Encoders for Object Detection from Point Clouds, IEEE CVPR 2019

[41] Zhou J et al., FVNet: 3D Front-View Proposal Generation for Real-Time Object Detection from Point Cloud, arXiv 1903.10750, 2019

[42] Shi S et al., Part-A^2 Net: 3D Part-Aware and Aggregation Neural Network for Object Detection from Point Cloud, arXiv 1907.03670, 2019

[43] Wang B, An J, and Cao J, Voxel-FPN: multi-scale voxel feature aggregation in 3D object detection from point clouds, arXiv 1907.05286, 2019

[44] Yang Z et al., STD: Sparse-to-Dense 3D Object Detector for Point Cloud, arXiv 1907.10471,2019

[45] Chen Y et al., Fast Point RCNN, arXiv 1908.02990, 2019

[46] Ngiam J et al., StarNet: Targeted Computation for Object Detection in Point Clouds, arXiv 1908.11069, 2019
[47] Zhu B et al., Class-balanced Grouping and Sampling for Point Cloud 3D Object Detection, arXiv 1908.09492, 2019
[48] Qi C et al., Deep Hough Voting for 3D Object Detection in Point Clouds, arXiv 1904.09664, 2019
[49] Meyer G et al., LaserNet: An Efficient Probabilistic 3D Object Detector for Autonomous Driving, arXiv 1903.08701, 2019
[50] Chen X et al., Monocular 3D Object Detection for Autonomous Driving, IEEE CVPR 2016
[51] F Chabot et al., Deep MANTA: A Coarse-to-fine Many-Task Network for joint 2D and 3D vehicle analysis from monocular image， IEEE CVPR， 2017
[52] A Mousavian et al., 3D Bounding Box Estimation Using Deep Learning and Geometry， IEEE CVPR， 2017
[53] H-N Hu et al., Joint Monocular 3D Vehicle Detection and Tracking， arXiv 1811.10742, 11， 2018
[54] T Roddick, A Kendall, R Cipolla, Orthographic Feature Transform for Monocular 3D Object Detection， arXiv 1811.08188, 11， 2018
[55] B Xu, Z Chen， Multi-Level Fusion based 3D Object Detection from Monocular Images, IEEE CVPR, 2018
[56] Z Qin， J Wang， Y Lu， MonoGRNet: A Geometric Reasoning Network for Monocular 3D Object Localization, arXiv 1811.10247, 11, 2018
[57] T He, S Soatto， Mono3D++: Monocular 3D Vehicle Detection with Two-Scale 3D Hypotheses and Task Priors, arXiv 1901.03446, 1, 2019
[58] Weng X, Kitani K, Monocular 3D Object Detection with Pseudo-LiDAR Point Cloud, arXiv 1903.09847, 2019
[59] Li B et al., GS3D: An Efficient 3D Object Detection Framework for Autonomous Driving, arXiv 1903.10955, 2019
[60] Ma X et al., Accurate Monocular Object Detection via Color-Embedded 3D Reconstruction for Autonomous Driving, arXiv 1903.11444, 2019
[61] Ku J, Pon A D, Waslander S L, Monocular 3D Object Detection Leveraging Accurate Proposals and Shape Reconstruction, arXiv 1904.01690, 2019
[62] Barabanau I et al., Monocular 3D Object Detection via Geometric Reasoning on Keypoints, arXiv 1905.05618, 2019
[63] Naiden A et al., Shift R-CNN: Deep Monocular 3d Object Detection with Closed-Form Geometric Constraints, arXiv 1905.09970, 2019
[64] Simonelli A et al., Disentangling Monocular 3D Object Detection, arXiv 1905.12365, 2019
[65] Jörgensen E, Zach C, Kahl F. Monocular 3D Object Detection and Box Fitting Trained End-to-End Using Intersection-over-Union Loss, arXiv 1906.08070, 2019
[66] Brazil G, Liu X. M3D-RPN: Monocular 3D Region Proposal Network for Object Detection， arXiv 1907.06038, 2019
[67] Wang X et al., Task-Aware Monocular Depth Estimation for 3D Object Detection, arXiv 1909.07701, 2019
[68] Marie J et al., Refined MPL: Refined Monocular PseudoLiDAR for 3D Object Detection in Autonomous Driving, arXiv 1911.09712, 2019
[69] Simonelli1 A et al., Towards Generalization Across Depth for Monocular 3D Object Detection, arXiv 1912.08035, 2020
[70] Li P et al., RTM3D: Real-time Monocular 3D Detection from Object Keypoints for Autonomous Driving, arXiv 2001.03343, 2020
[71] Cai Y et al., Monocular 3D Object Detection with Decoupled Structured Polygon Estimation and Height-Guided Depth Estimation, arXiv 2002.01619, 2020
[72] Li P, Chen X, Shen S, Stereo R-CNN based 3D Object Detection for Autonomous Driving， arXiv 1902.09738, 2019
[73] Qin Z, Wang J, Lu Y, Triangulation Learning Network: from Monocular to Stereo 3D Object Detection， arXiv 1906.01193, 2019
[74] Wang Y et al., Pseudo-LiDAR from Visual Depth Estimation: Bridging the Gap in 3D Object Detection for Autonomous Driving, arXiv 1812.07179, 2019
[75] Pon A et al., Object-Centric Stereo Matching for 3D Object Detection, arXiv 1909.07566, 2019
[76] Danzer A et al., 2D Car Detection in Radar Data with PointNets, arXiv 1904.08414, 2019
[77] Nabati R, Qi H, RRPN: Radar Region Proposal Network for Object Detection in Autonomous Vehicles, arXiv 1905.00526, 2019
[78] Major B et al., Vehicle Detection with Automotive Radar Using Deep Learning on Range-Azimuth-Doppler Tensors, IEEE ICCV workshop, 2019
[79] D Eigen, C Puhrsch and R Fergus, Depth Map Prediction from a Single Image using a Multi-Scale Deep Network, arXiv 1406.2283, 2014
[80] D Eigen, R Fergus, Predicting Depth, Surface Normals and Semantic Labels with a Common Multi-Scale Convolutional Architecture, IEEE ICCV 2015
[81] P Wang et al., Towards Unified Depth and Semantic Prediction From a Single Image, IEEE CVPR 2015
[82] F Liu et al., Deep Convolutional Neural Fields for Depth Estimation from a Single Image, IEEE CVPR 2015
[83] R Garg et al., Unsupervised CNN for Single View Depth Estimation: Geometry to the Rescue, ECCV 2016
[84] C Godard et al., Unsupervised Monocular Depth Estimation with Left-Right Consistency, IEEE CVPR 2017
[85] Y Kuznietsov et al., Semi-Supervised Deep Learning for Monocular Depth Map Prediction, IEEE CVPR 2017
[86] T Zhou et al., Unsupervised Learning of Depth and Ego-Motion from Video, IEEE CVPR 2017
[87] Y Luo et al., Single View Stereo Matching, arXiv 1803.02612, 2018
[88] J Jiao et al., Look Deeper into Depth: Monocular Depth Estimation with Semantic Booster and Attention-Driven Loss, ECCV, 2018
[89] C Wang et al., Learning Depth from Monocular Videos using Direct Methods, IEEE CVPR 2018
[90] V Casser et al., Depth Prediction Without the Sensors: Leveraging Structure for Unsupervised Learning from Monocular Videos, arXiv 1811.06152, 2018
[91] X Qi et al., GeoNet: Geometric Neural Network for Joint Depth and Surface Normal Estimation, IEEE CVPR 2018
[92] Z Yin, J Shi, GeoNet - Unsupervised Learning of Dense Depth, Optical Flow and Camera Pose, IEEE CVPR 2018
[93] Zou Y, Luo Z, Huang J., DF-Net: Unsupervised Joint Learning of Depth and Flow using Cross-Task Consistency, arXiv 1809.01649, 2018
[94] Z Yang et al., LEGO: Learning Edge with Geometry all at Once by Watching Videos, *AAAI*, 2019
[95] Ranjan A et al., Competitive Collaboration: Joint Unsupervised Learning of Depth, Camera Motion, Optical Flow and Motion Segmentation, arXiv 1805.09806, 2019
[96] Casser V et al., Unsupervised Monocular Depth and Ego-motion Learning with Structure and Semantics, arXiv 1906.05717, 2019
[97] Bian J et al., Unsupervised Scale-consistent Depth and Ego-motion Learning from Monocular Video, arXiv 1908.10553, 2019
[98] Andraghetti L et al., Enhancing self-supervised monocular depth estimation with traditional visual odometry, arXiv 1908.03127, 2019
[99] Zhou L, Fang J, Liu G, Unsupervised Video Depth Estimation Based on Ego-motion and Disparity Consensus, arXiv 1909.01028, 2019
[100] Shi Y et al., Self-Supervised Learning of Depth and Ego-motion with Differentiable Bundle Adjustment, arXiv 1909.13163, 2019
[101] Wang G et al., Unsupervised Learning of Depth, Optical Flow and Pose with Occlusion from 3D Geometry, arXiv 2003.00766, 2020
[102] A Kendall et al.， End-to-End Learning of Geometry and Context for Deep Stereo Regression, ICCV, 2017
[103] C Zhou et al., Unsupervised Learning of Stereo Matching, ICCV 2017
[104] J R Chang, Y S Chen, Pyramid Stereo Matching Network, arXiv 1803.08669, 2018
[105] G Yang et al., SegStereo: Exploiting Semantic Information for Disparity, ECCV 2018.
[106] J Zhang et al., DispSegNet: Leveraging Semantics for End-to-End Learning of Disparity Estimation from Stereo Imagery, arXiv 1809.04734, 2019

[107]Jonas Uhrig et al., Sparsity Invariant CNNs，*Int. Conf. on 3D Vision*，8，2017

[108]A Eldesokey et al., Propagating Confidences through CNNs for Sparse Data Regression，BMCV，5，2018

[109]HMS-Net: Hierarchical Multi-scale Sparsity-invariant Network for Sparse Depth Completion, arXiv 1808.08685, 2018

[110]S Minaee et al., Image Segmentation Using Deep Learning: A Survey, arXiv 2001.05566, April 2020

[111]G Ciaparrone et al., Deep Learning in Video Multi-Object Tracking: A Survey, arXiv 1907.12740, Nov. 2019

[112]Nick Schneider et. Al., RegNet: Multimodal Sensor Registration Using Deep Neural Networks, arXiv 1707.03167, 2017.

[113]G Iyer et al., CalibNet: Self-Supervised Extrinsic Calibration using 3D Spatial Transformer Networks, arXiv 1803.08181, 2018

[114]F Ma，S Karaman，Sparse-to-Dense: Depth Prediction from Sparse Depth Samples and a Single Image，arXiv 1709.07492, 9 2017

[115]Y Zhang, T Funkhouser，Deep Depth Completion of a RGB-D Image，IEEE CVPR，2018

[116]Z Chen et al., Estimating Depth from RGB and Sparse Sensing，ECCV，4，2018

[117]K Park, S Kim, K Sohn，High-precision Depth Estimation with the 3D LiDAR and Stereo Fusion，IEEE ICRA，5，2018

[118]F Ma et al., Self-Supervised Sparse-to-Dense: Depth Completion from LiDAR and Mono Camera，arXiv 1807.00275，7，2018

[119]M Jaritz et al., Sparse and Dense Data with CNNs: Depth Completion and Semantic Segmentation，arXiv 1808.00769，8，2018

[120]M Dimitrievski, P Veelaert, W Philips，Learn Morphological Operators for Depth Completion, *Int. Conf. on Advanced Concepts for Intelligent Vision Systems*, 7, 2018

[121]J Qiu et al., DeepLiDAR: Deep Surface Normal Guided Depth Prediction from LiDAR and Color Image, arXiv 1812.00488, 12, 2018

[122]Y Yang, A Wong, S Soatto，Dense Depth Posterior (DDP) from Single Image and Sparse Range, arXiv 1901.10034, 1, 2019

[123]S. Shivakumar et al., DFuseNet: Fusion of RGB and Sparse Depth for Image Guided Dense Depth Completion，arXiv 1902.00761，2，2019

[124]Gansbeke W V et al, Sparse and noisy LiDAR completion with RGB guidance and uncertainty, arXiv 1902.05356, 2019

[125]T-H Wang et al., 3D LiDAR and Stereo Fusion using Stereo Matching Network with Conditional Cost Volume Normalization, arXiv 1904.02917, 2019

[126]Zhong Y et al., Deep RGB-D Canonical Correlation Analysis For Sparse Depth Completion, arXiv 1906.08967, 2019

[127]Eldesokey A, Felsberg M, Khan F S, Confidence Propagation through CNNs for Guided Sparse Depth Regression, arXiv 1811.01791, 2019

[128]Tang J et al., Learning Guided Convolutional Network for Depth Completion, arXiv 1908.01238, 2019

[129]Zhang Y et al., DFineNet: Ego-Motion Estimation and Depth Refinement from Sparse, Noisy Depth Input with RGB Guidance, arXiv 1903.06397, 2019

[130]Liu H et al., PLIN: A Network for Pseudo-LiDAR Point Cloud Interpolation, arXiv 1909.07137, 2019

[131]Xu Y et al., Depth Completion from Sparse LiDAR Data with Depth-Normal Constraints, arXiv 1910.06727, 2019

[132]X Chen et al., Multi-View 3D Object Detection Network for Autonomous Driving，arXiv 1611.07759, 2017.

[133]D Matti, H Kemal Ekenel, J-P Thiran, Combining LiDAR Space Clustering and Convolutional Neural Networks for Pedestrian Detection, arXiv 1710.06160, 2017

[134]Z Wang et al., Fusing Bird's Eye View LIDAR Point Cloud and Front View Camera Image for Deep Object Detection, arXiv 1711.06703, 2018

[135]D Xu et al., PointFusion: Deep Sensor Fusion for 3D Bounding Box Estimation, arXiv 1711.10871, 2018.

[136]X Du et al., A General Pipeline for 3D Detection of Vehicles，arXiv 1803.00387, 2018

[137]K Shin et al., RoarNet: A Robust 3D Object Detection based on RegiOn Approximation Refinement, arXiv 1811.03818, 2018

[138]J Ku et al., Joint 3D Proposal Generation and Object Detection from View Aggregation, arXiv 1712.02294, 2018

[139]C Qi et al., Frustum PointNets for 3D Object Detection from RGB-D Data, arXiv 1711.08488, 2018.

[140]M Liang et al., Deep Continuous Fusion for Multi-Sensor 3D Object Detection, ECCV, 2018

[141]D Frossard, R Urtasun, End-to-end Learning of Multi-sensor 3D Tracking by Detection, arXiv 1806.11534, 2018.

[142]Zhao X et al., 3D Object Detection Using Scale Invariant and Feature Reweighting Networks, arXiv 1901.02237, 2019

[143]Sindagi V A, Zhou Y, Tuzel O, MVX-Net: Multimodal VoxelNet for 3D Object Detection, arXiv 1904.01649, 2019

[144]Meyer G et al., Sensor Fusion for Joint 3D Object Detection and Semantic Segmentation, arXiv 1904.11466, 2019

[145]Ku J et al., Improving 3D Object Detection for Pedestrians with Virtual Multi-View Synthesis Orientation Estimation, arXiv 1907.06777, 2019

[146]Raffiee A, Irshad H, Class-specific Anchoring Proposal for 3D Object Recognition in LIDAR and RGB Images, arXiv 1907.09081, 2019

[147]You Y et al., Pseudo-LiDAR++: Accurate Depth for 3D Object Detection in Autonomous Driving, arXiv 1906.06310, 2019

[148]Deng J, Czarnecki K, MLOD: A multi-view 3D object detection based on robust feature fusion method, arXiv 1909.04163, 2019

[149]Li P, Liu S, Shen S, Multi-Sensor 3D Object Box Refinement for Autonomous Driving, arXiv 1909.04942, 2019

[150]Cadena C et al., Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age, arXiv 1606.05830, 2016

[151]S Milz et al., Visual SLAM for Automated Driving: Exploring the Applications of Deep Learning, *IEEE CVPR Workshop*, 2018

[152]B Huang, J Zhao, J Liu, A Survey of Simultaneous Localization and Mapping, arXiv 1909.05214, Aug. 2019

[153]Alahi A et al., Social LSTM: Human Trajectory Prediction in Crowded Spaces, IEEE CVPR 2016

[154]Pfeiffer M et al., A Data-driven Model for Interaction-aware Pedestrian Motion Prediction in Object Cluttered Environments, arXiv 1709.08528, 2018

[155]Gupta A et al., Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks, CVPR 2018

[156]Sedegian A et al., SoPhie: An Attentive GAN for Predicting Paths Compliant to Social and Physical Constraints, arXiv 1806.01482, Sep. 2018

[157]Vemula A, Muelling K, Oh J, Social Attention: Modeling Attention in Human Crowds, ICRA, 2018

[158]Hoy M et al., Learning to Predict Pedestrian Intention via Variational Tracking Networks, *IEEE ITSC*, Nov. 2018

[159]Xue H, Huynh D, Reynolds M, Location-Velocity Attention for Pedestrian Trajectory Prediction, *IEEE WACV*, 2019

[160]Haddad S et al., Situation-Aware Pedestrian Trajectory Prediction with Spatio-Temporal Attention Model, *IEEE WACV*, 2019

[161]Zhang P et al., SR-LSTM: State Refinement for LSTM towards Pedestrian Trajectory Prediction, arXiv 1903.02793, Mar. 2019

[162]Zhu Y. et al., StarNet: Pedestrian Trajectory Prediction using Deep Neural Network in Star Topology, arXiv 1906.01797, June 2019

[163]Zhao T et al., Multi-Agent Tensor Fusion for Contextual Trajectory Prediction, arXiv 1904.04776, July 2019

[164]Amirian J, Hayet J-B, Pettre J, Social Ways: Learning Multi-Modal Distributions of Pedestrian Trajectories with GANs, *IEEE CVPR workshop*, 2019

[165]Li Y., Which Way Are You Going? Imitative Decision Learning for Path Forecasting in Dynamic Scenes, IEEE CVPR 2019

[166]Chandra R et al., TraPHic: Trajectory Prediction in Dense and Heterogeneous Traffic Using Weighted Interactions, IEEE CVPR 2019

[167] Choi C, Dariush B, Learning to Infer Relations for Future Trajectory Forecast, IEEE ICCV 2019
[168] Liang J et al., Peeking into the Future: Predicting Future Person Activities and Locations in Videos, IEEE CVPR 2019
[169] Fan H et al., An Auto-tuning Framework for Autonomous Vehicles, arXiv 1808.04913, 2018
[170] Wolf P et al., Adaptive Behavior Generation for Autonomous Driving using Deep Reinforcement Learning with Compact Semantic States, arXiv1809.03214, 2018
[171] Sun L, Zhan W, Tomizuka M, Probabilistic Prediction of Interactive Driving Behavior via Hierarchical Inverse Reinforcement Learning, arXiv 1809.02926, 2018
[172] Xu Z, Tang C, Tomizuka M, Zero-shot Deep Reinforcement Learning Driving Policy Transfer for Autonomous Vehicles based on Robust Control, arXiv 1812.03216, 2018
[173] Cui H et al., Multimodal Trajectory Predictions for Autonomous Driving using Deep Convolutional Networks, arXiv 1809.10732, 2018
[174] Ma Y et al., TrafficPredict: Trajectory Prediction for Heterogeneous Traffic-Agents, arXiv 1811.02146, 2018
[175] Bansal M, Krizhevsky A, Ogale A, ChauffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst, arXiv 1812.03079, 2018
[176] Codevilla F et al., Exploring the Limitations of Behavior Cloning for Autonomous Driving, arXiv 1904.08980, 2019
[177] Hoel C-J et al., Combining Planning and Deep Reinforcement Learning in Tactical Decision Making for Autonomous Driving, arXiv 1905.02680, 2019
[178] Moghadam M, Elkaim G H, A Hierarchical Architecture for Sequential Decision-Making in Autonomous Driving using Deep Reinforcement Learning, 1906.08464, 2019
[179] Lee N et al., DESIRE: Distant Future Prediction in Dynamic Scenes with Interacting Agents, arXiv 1704.04394, 2017
[180] Srikanth S et al., INFER: INtermediate representations for FuturE pRediction, arXiv 1903.10641, 2019
[181] Si W, Wei T, Liu C, AGen: Adaptable Generative Prediction Networks for Autonomous Driving, *IEEE IV*, 2019
[182] Li J, Ma H, Tomizuka M, Conditional Generative Neural System for Probabilistic Trajectory Prediction, arXiv 1905.01631, 2019
[183] Li J, Ma H, Zhan W, Tomizuka M, Coordination and Trajectory Prediction for Vehicle Interactions via Bayesian Generative Modeling, arXiv 1905.00587, 2019
[184] Li J, Ma H, Tomizuka M, Interaction-aware Multi-agent Tracking and Probabilistic Behavior Prediction via Adversarial Learning, arXiv 1904.02390, 2019
[185] Hu Y, Sun L, Tomizuka M, Generic Prediction Architecture Considering both Rational and Irrational Driving Behaviors, arXiv 1907.10170, 2019
[186] Cho K et al., Deep Predictive Autonomous Driving Using Multi-Agent Joint Trajectory Prediction and Traffic Rules, IEEE IROS, 2019
[187] Li X, Ying X, Chuah M C, GRIP: Graph-based Interaction-aware Trajectory Prediction, arXiv 1907.07792, 2019
[188] Chai Y et al., Multi-Path: Multiple Probabilistic Anchor Trajectory Hypotheses for Behavior Prediction, arXiv 1910.05449, 2019
[189] Lee D et al., Joint Interaction and Trajectory Prediction for Autonomous Driving using Graph Neural Networks, arXiv 1912.07882, 2019
[190] Chandra R et al., Forecasting Trajectory and Behavior of Road-Agents Using Spectral Clustering in Graph-LSTMs, arXiv 1912.01118, 2019
[191] Li J et al., Social-WaGDAT: Interaction-aware Trajectory Prediction via Wasserstein Graph Double-Attention Network, arXiv 2002.06241, 2020
[192] Li J et al., EvolveGraph: Heterogeneous Multi-Agent Multi-Modal Trajectory Prediction with Evolving Interaction Graphs, arXiv 2003.13924, 2020
[193] Hu Y, Zhan W, Tomizuka M, Scenario-Transferable Semantic Graph Reasoning for Interaction-Aware Probabilistic Prediction, arXiv 2004.03053, 2020
[194] Gao J et al., VectorNet: Encoding HD Maps and Agent Dynamics from Vectorized Representation, arXiv 2005.04259, 2020
[195] S Mozaffari et al., Deep Learning-based Vehicle Behavior Prediction for Autonomous Driving Applications: A Review, arXiv 1912.11676, Dec. 2019.
[196] S Kuutti et al., A Survey of Deep Learning Applications to Autonomous Vehicle Control, arXiv 1912.10773, Dec. 2019
[197] M Bojarski et al., End to End Learning for Self-Driving Cars, arXiv 1604.07316, 2016
[198] E Santana, G Hotz, Learning a Driving Simulator, arXiv 1608.01230, 2016
[199] Y Tian et al., DeepTest: Auto Testing of DNN-driven Autonomous Cars, arXiv 1708.08559， 2017
[200] S Hecker, D Dai, L Gool, Failure Prediction for Autonomous Driving, arXiv 1805.01811， 2018
[201] Y Chen et al., LiDAR-Video Driving Dataset: Learning Driving Policies Effectively, IEEE CVPR 2018
[202] S Hecker, D Dai, L Gool, End-to-End Learning of Driving Models with Surround-View Cameras and Route Planners, ECCV 2018
[203] Grigorescu S et al., NeuroTrajectory: A Neuroevolutionary Approach to Local State Trajectory Learning for Autonomous Vehicles, arXiv 1906.10971, 2019
[204] T Wheeler et al., Deep Stochastic Radar Models, IEEE IV, 2017
[205] H Alhaija et al., Augmented Reality Meets Computer Vision: Efficient Data Generation for Urban Driving Scenes, arXiv 1708.01566， 2017
[206] VSR Veeravasarapu, C Rothkopf, R Visvanathan, Adversarially Tuned Scene Generation, arXiv 1701.00405, 2017
[207] X Yue et al， A LiDAR Point Cloud Generator: from a Virtual World to Autonomous Driving, arXiv 1804.00103, 2018
[208] K Elmadawi et al., End-to-end sensor modeling for LiDAR Point Cloud, arXiv 1907.07748v1，2019，7
[209] R McAllister et al. Concrete Problems for Autonomous Vehicle Safety: Advantages of Bayesian Deep Learning，*IJCAI*，2017
[210] T Dreossi et al., VERIFAI: A Toolkit for the Formal Design and Analysis of Artificial Intelligence-Based Systems, *Int. Conf. on Computer Aided Verification* (CAV), 2019
[211] Marcus G, Deep Learning: A Critical Appraisal, arXiv 1801.00631, 2018
[212] Yullie A, Liu C, Deep Nets: What have they ever done for Vision? arXiv 1805.04025, 2019
[213] Thys S, Ranst W, Goedeme T, Fooling Automated Surveillance Cameras: Adversarial Patches to Attack Person Detection, arXiv 1904.08653, 2019
[214] Jia Y et al., Fooling Detection Alone Is Not Enough: Adversarial Attack Against Multiple Object Tracking, ICLR, 2020
[215] KITTI data: http://www.cvlibs.net/datasets/kitti/raw_data.php
[216] BDD: http://bdd-data.berkeley.edu/.
[217] Baidu ApolloScope: http://apolloscape.auto/
[218] NuScenes: https://www.nuscenes.org/
[219] Udacity: https://github.com/udacity/self-driving-car/tree/master/datasets
[220] Ford ArgoVerse: https://www.argoverse.org/data.html
[221] Google WayMo: https://waymo.com/open
[222] Lyft Level 5: https://level5.lyft.com/dataset/
[223] HONDA dataset: https://usa.honda-ri.com/H3D
[224] Scale and Hesai's PandaSet: https://scale.com/open-datasets/pandaset
[225] NGSIM data: https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm
[226] HighD dataset: https://www.highd-dataset.com/
[227] INTERACTION dataset: https://interaction-dataset.com/