Neural Network Architectures for Lane Detection and Road Segmentation in Self-Driving Cars

By Dr. George Tsoulos

Professor of Electrical and Computer Engineering, Cyprus University of Technology (CUT)

1. Introduction

Road and lane detection are essential components for the environmental perception of Advanced Driver Assistance Systems (ADAS) and self-driving cars (SDC) [1]. It predicts the map specifications that will be used for localization and planning algorithms. As well, it is critical when it comes to path planning in ADAS models. While there are many well-developed methods for lane detection and segmentation, the most important challenge is the generalization problem related to the dataset differences in real-world applications. Classes that are not found in the pre-trained model, ad-hoc tuning instead of training on real-world instances, limited generalization potential due to the small batch sizes, handcrafted feature and synthetic data usage, ad-hoc region proposal methods will lead to a non-robust network on real-world application data even though the segmentation network in question could be successful at the image dataset [2]. This is why networks overfit to dataset related to the problem of encoder-decoder architectures. However, a few studies have been done to explicitly investigate the generalization potential of encoder-decoder architectures. There are only several quantification reports with large generalization benchmark dataset analysis to advantage encoder-decoder architectures in the area of Object Detection and Segmentation.

Similar to human-made maps, lane detection is a necessary component of the environmental perception module in Advanced Driver Assistance Systems (ADAS) and self-driving car (SDC) technologies. The lane detection tool identifies the vehicle's position in the traffic lane and plays a crucial role in Autonomous Vehicle (AV) trajectory planning [3]. Different pre-trained networks have been studied for AV lane detection, refining them in lane-annotation fashion for better segmentation. However, it is observed that the small batch size limits the generalization potential of neural networks during training, a significant challenge in datasets with high variations in annotations and real-world tasks. The state-of-the-art networks are widely overtrained when ad-hoc tuning is conducted on pre-trained networks. Moreover, hand-crafted features and synthetic data usage techniques, like data augmentation, boost performances on adapted networks chosen from pre-developed networks. A large number of images that can be captured from the real-world environment make data consistency with the real-world orientation and scale a critical necessity in the lane detection datasets. The combination of pre-trained neural networks and data consistency boosting techniques lacks explicit methods among the state-of-the-art lanes detection methods and quantitative evaluations to provide generalization performance.

1.1. Background and Motivation

In this article, a detailed survey of modern techniques for lane detection has been presented and classified into five corresponding categories. The current status of existing approaches in the literature is compared and summarized with respect to the several performance measures. The associated challenges of each category have further been explored and discussed in a comprehensive manner. The challenges associated with existing methods have also been thoroughly discussed in this work. Lastly, future potential directions have also been identified to serve the related domain.

[3] [4] [5]In recent years, deep learning (DL) techniques have been successfully applied to many areas of information technology. The application of methods of this branch, called artificial intelligence (AI) methods, allow successful solutions to the problem of self-driving cars, so allowing drivers to rely on vehicles more and more often. Active development towards fully autonomous cars means development of these methods. A great challenge here is lane detection. It concerns achieving high precision while, in the case of road segmentation, also achieving an appropriate generalization in different conditions, especially different weather conditions. Simultaneously, the real-time character of the methods is fundamental due to the real-time character of the task and the efficiency of cars. Despite applications have been developed, the problem is still current and challenging. There are many methods addressing turnover,,. This review can be helpful for the formulation of new challenges.

1.2. Research Objectives

Lane detection and vehicle detection play a very important role in unmanned driving. Simultaneous vehicle and lane detection in a vehicle-following scene can effectively deal with complex traffic scenarios while decreasing the calculation burden. In this research, a unified deep learning-based method is presented. The core technology of detecting the lane and vehicle is deep learning feature extraction. Transfer learning is being utilized under limited data collection conditions. The number of model parameters is reduced by using the MobileNetV3 feature extraction network [6].

Lane and road segmentation are complex problems in computer vision that have seen the application of a variety of modern convolutional neural network architectures. Recent efforts have developed approaches with progressively more lightweight neural networks in an effort to cater for embedded systems as well as any platforms requiring real-time or low-latency performance. This highlights the fact that current applications and challenges faced within the fields of lane and road segmentation lie at the intersection of common sub-fields of computer vision and the specific requirements of designing and training networks for in-vehicle applications as well as systems with embedded constraints. Both lane and road segmentation are usually treated in isolation. It is beneficial to approach them in combination, because other classes of bias wherein one class is more prominent can be overcome by having multiple response targets. This pairs with the underlying theme of this publication which is to reduce the overall complexity of the solution space by decoupling structures from the object classes being detected [2].

2. Fundamentals of Lane Detection and Road Segmentation

Vision-based methods for lane detection include edge extraction and Hough transform. Convolutional neural networks have also shown state-of-the-art performance, freeing the need for manual feature design. There are also some CNN-based approaches such as Mask-RCNN and U-Net. Numbers of researchers proposed multi-task networks that would output lane or/and road detection results by feeding in images directly. Fully convolutional networks were also popularly used for lane or road detection. Instance-segmentation networks (AMED and LaneATT) had even achieved the state-of-the-art performance for road mapping and lane marking detection. However, those networks are designed for road mapping or lane marking detection and would not consider the connections between roads and lanes. In general, most existing lane detection models used only lane pixel-level information for lane detection.

There are two identical reasons for the given list of problems in lane detection and road segmentation. Lane detection aims at detecting lane markers' positions based on retinal images and has been widely studied over the past few decades [7]. Detecting road regions has received more attention due to the growing significance of autonomous driving.

High-precision and real-time perception of the road environment is the key aspect of advanced driving assistance systems (ADAS) and automatic driving (AD) [8]. Developing a clean and reliable understanding of the road environment is both necessary and irrelevant. Recognizing roads including sidewalks, marked lanes, lane boundaries and intersections is essential for planning and control of vehicle motion. Detecting lanes is a focal problem in vehicle perception, while road segmentation is periodic information reduced.

2.1. Basic Concepts in Image Processing

In the study by Joon You Yang and his team in the paper "Fast and Accurate Lane Detection via Graph Structure and Disentangled Representation Learning", the authors presents a novel method for detecting Lane and Road in automotive cars by leveraging the underlying structure of image representations through Graphs, and propose a novel method to learn the representation in a disentangled manner so as to make the network focus better on understanding different objects, lanes, and roads in our case [2]. The balance between centralized and disentangled representations and downstream tasks of representing camera data is automatically optimized, providing up to 15% improvement in setting the optimal hyper parameters and reduces the computation cost using GCN, SEM, and lane segmentation block in the backbone of our framework.

[9] Lane and road detection have made significant contributions to the development of AI-enhanced maps, road-facing perceptio and autonomous vehicles in the automobile domain [10]. Among all the input channels, visual input such as dashboard cameras has been widely used in previous research for practical reasons like being cost effective and requiring less encoding, transmission, and decoding latency while providing sufficient information for Safe HMI, DAS, and ADAS for the vehicles.

2.2. Challenges and Importance in Self-Driving Cars

In the literature, a large number of approaches can be found in the context of this study 20, 33, 40, 41, 42. The general structure of the lane detection problem can be divided into two categories: traditional methods and deep learning-based methods. Traditional methods use hand-crafted features and shallow model learning, which may lead to unstable performance. To address this challenge, deep learning-based lane detection methods propose trainable end-to-end with no need for hand-crafted feature designing [11]. As well known deep learning can achieve better performance than shallow learning in terms of different scale, experiment and perspective reasons, So, deep learning is largely involved for lane detection research recent years.

Lane detection and road segmentation have a great importance in self-driving cars [12]. They are fundamental for building high-level scenes or maps, supporting the robots planning and decision-making in the system. However, with day-to-day environment changes on the roads, these systems are faulty. For example, in challenging scenarios like poor lighting, low light, light sticking, vehicle tip over junction lines, dashed lines and their occlusion problems with shadows, noise or even the road suddenly becomes difficult. Therefore, achieving accurate and robust lane detection in real time domain is an important issue [10]. Unintended lane departure is a leading cause of motor vehicle collisions and highway vehicles collisions, highlighting the importance of lane detection in self-driving cars. Inaccurate lane detection or not detecting the lane are defined as mistakes made by the tarmac markings detection module that are the main components of the lane detection.

3. Neural Networks in Computer Vision

This section provides brief introductions to assist the understanding of the architectures used for lane detection and road segmentation. Classifiers and segmenters alike tend to employ neural networks. They are manually trained and use various models that consume input data from images or point clouds. The architectures have evolved with each passing paper, and the final models were often effective and efficient in tailoring data to the problem of lane and road boundary detection and segmentation. Some deep architectures have managed to out-esteem current state-of-the-art solutions [2].

Lane detection and road segmentation are essential tasks in the advancement of intelligent transportation systems such as self-driving cars. Matthias Dang et al. [13] pinpoint that lane detection has seen a recent decrease in popularity, especially in contrast to vehicle detection. A plethora of different computer vision techniques are employed for support in data collection and analysis, including mathematical-, machine learning-, and neural network-based methods. Various papers also discuss the merits of deep neural networks for lane detection, outlining architectures that have improved upon and challenged current state-of-the-art solutions.

3.1. Convolutional Neural Networks (CNNs)

The proposed solution TwinLiteNetPlus efficiently performs real-time joint segmentation of multiple classes with a single network. The use of multi-resolution attention gates yields an enhanced learning capacity for the model, contributing to its successful multi-task learning. By incorporating auxiliary decoder prediction modules, loss re-weighting, and depth-wise separable convolutions, the architecture proves its effectiveness in its ability to identify and parameterizes road structures like lanes and drivable areas. Many state-of-the-art deep learning-based lane detection models include prediction modules for straight lanes, lanes involving curvatures and lanes exhibiting discontinuities. Lateral shifts to learn vertical affine transformations of features enhance the generalization capacity of the model to newer data distributions. The utilization of a Dual-Dilated Spatial Pyramid Atrous Network (D-TwinNet) generalizes the feature processing to lead the model to learn a set of disentangled, scale-specific feature maps that help it generate a hybrid refinement of predictions from multiple pyramid resolutions.

Convolutional Neural Networks (CNNs) [14] are an effective tool for image classification, object detection, and image segmentation tasks. CNNs were initially deployed to solve the task of handwritten digit recognition, where the performance gradually increased. By 2012, SuperVision with a high-resolution representation by the classifier (AlexNet) had narrowly won the ImageNet competition. Convolutional layers have helped CNNs recognise precise patterns. New sets of filters can be learned for each convolutional layer. The ReLU activation function helps solve the problem of vanishing gradients. Pooling layers spatially shrink the network into progressively smaller dimensions. Fully connected layers link all neurons in the previous layer to the current layer. The combination of convolutional, activation, pooling, and fully connected layers helps CNNs recognise spatial local features and generate different feature sets. CNNs can leverage this ability to retain spatial information, to solve pixel-wise labelling problems. This makes CNNs well suited for lane detection and road segmentation tasks [12].

3.2. Recurrent Neural Networks (RNNs)

In, a new way of fully end-to-end cascaded convolutional-recurrent network, which forecasts the localization of a robot's forward driving path at arbitrary horizons was proposed, which is conditioned on current RGB sequence data. We are persuaded that < > can be effectively used to solve two tasks that are generally carried out in separate ways in an interconnected way-through LTL formulation and share the contextual, temporal and spatial information at the same time. There are several researches in which a four-pathway Le-Net was applied by depthwise separable convolution to greatly reduce the models' number of parameters, computation loads and also stabilize the model's rapid training. A modified variants of PathagNet and then a residual network based structure with PyramidNet's idea have been proposed on lateral marking networks that boost the lane segmentation speeds from the low prevalence of the original PathagNet's fully convolutional network that require about 45 ms per frame on Titan Xp.caffold. Without doubt, the trained network had a pedging size can let the intersection of itself and largest of ground truth be larger by following strategies: it replaces a 3x3 full convolutional layer with the 1x1 depthwise separable convolutional layer, applys depthwise separable convolutional layer in skip sequences with 3x3 spatial kernel, and pads the model with a bigger receptive fiels.

Recurrent Neural Networks (RNNs) are one of the deep learning architectures that excel at pixel-level classification tasks and, unlike convolutional networks, can explicitly model regularity in spatio-temporal data [12]. Traditional methods in this field are mainly based on the extraction of hand-crafted features from artificial images, and/or on sliding window classifiers, which are much less effective in detecting the lateral markings of the road. Other types of misalignments of the convolution operation and pedestrian crossing treatments lead to difficulties of detection. RNN is also good at lazily processing input sequence context elements with extremely high energy and great potential for asynchronous tasks, and is robust to time index perturbation. Segmenting lateral lines directly with CNNs in a lane system is not sufficient for every lane shape, especially when lane segments are lacking. We combine the RNN with the lane system in the form of a depthwise separable convolution-based encoder, changing the segmentation standard from direct segmentation to a multi-stage box expansion path; experiments demonstrate the robustness and recognition ability of the detection of the proposed network in detecting various lane fragments that are irregularly shaped, being occluded and deformed [15].

3.3. Deep Learning Architectures

Experiments state that the SPLTNet model is more practical for lane detection in self-driving cars, because of reduced computational complexity and powerful performance in comparison to state-of-the art and state-of-the-practice architectures. This architecture can detect lanes irrespective of the various changes that appear in the frame in only one forward pass. SUMNet, despite its efficiency to fully capture the global contextual lane information, has difficulty predicting and locating the hidden lane occulting regions. So that accuracy was relatively reduced. Further lane detection is divided into feature extraction, point prediction, and lane prediction. PINet adds a branch to the high-level points detection to improve the performance of lane end-tunnel detection and proposes detailed modifications about network architecture and training strategy in order to enhance the end-to-end training. The FCD-PAC model also brought some performance improvements on challenging scenarios. However, both PINet and SUMNet can bring out inordinate redundancy between the two ends of branch predictions. [16] [7]

Various state-of-the-art architectures such as RONELD, D-shaped network, SegNet, VPG-Net and SPLTNet constitute the next generation of lane detection network as it is facilitating lane change detection with all advanced curve fitting algorithm. The RONELD model has been designated mainly for their various lane detection metrics and the tolerance of challenging environments. This model employs RNN for identifying the lanes in conditional frames in order to deal with the challenging exterior including shadows, occlusions, and crosswalks as well as to determine lanes in light traffic. Yet, RONELD failed to identify lanes in highways and tunnels and most importantly smaller roads due to heavy architectural overheads. Thereby, annotation efforts are significantly increasing to bridge the gap of each new designated scenarios. Similar to RONELD, FCN has advantages considering that it just requires the output of the previous layer in order to produce predictions for each pixel.

4. Lane Detection Techniques

Recently, lane detection methods are generally divided into two categories. End-to-end trainable models are used for steering angle output and most of the models are regression based; their outputs are commonly a vector of angles to the ground truth lane markers. These types of models are designed with different methods, such as drawing Bezier curves instead of lines to make it compatible with neural networks and CNN architecture with an RNN to predict lane lines as a spatio-temporal sequence. There are also methods that predict output as a vanishing line, decoding lanes from previously calculated central lane [14]. Some road segmentation methods can also use lane detection as a precursor step. Deep Learning methods also started to make use of road segmentation annealing methods and combining RGB and Depth information to predict road masks separately from image information.

Lane detection plays a crucial role in applications such as autonomous vehicles, lane following, and intelligent transportation systems [11]. Many neural network models have been proposed for lane detection, such as the "supports of CNN" and U-Net which are used in image-to-image predictions [7]. The outputs of these models can be processed using tools like Hough transform. Learning model weights in semantic segmentation has shown promising results, for example, the wait loss on the feature maps is minimized in the spatial CNN model for lane detection. U-Net and Variational autoencoder Variational Autoencoder (VAE) are modified for lane detection in VPG-Net which uses a regression-based model to predict the vanishing point. Other models can use a combination of CNN and RANSAC for lane detection.

4.1. Traditional Methods

On top of this, the traditional methods cannot effectively detect a vanishing point, which is a single coordinate in the original input image generated from the perspective relation of road, and vehicles itself. In a vanishing point-driven scenario, traditional methods just focus on horizontal direction information, meaning they lose the depth information. Therefore, traditional methods show high stability by detecting them as one of the targets in images. The traditional methods ignore the structure of lane image and vanishing point. Notice that beyond finding the representative point of lane in image ends (sides of lane), the traditional methods pay the equal attention to each pixel in lane image. As a result, the traditional methods confuse out the lane elements among other structure information of lane image [17]. On the other hand, these traditional methods do not detect lanes well in curves and show performance drops [7].

Traditional lane detection relies on color-based, texture-based and intrinsic features such as edges, which fail to identify lanes with poor illumination, weather conditions, the number of lanes, or complicated ground textures. Additionally, the lane detection results are significantly affected by the search range of the region of interest, the slope or straightness of lanes, and exterior conditions. Besides, traditional lane paradigm detection methods have a series of problems, such as the complicated environment, limited sensitivity, and low accuracy. Therefore, traditional lane detection methods cannot meet the requirements of autonomous driving systems for lane detection [16]. In autonomous driving, traditional lane detection methods need to output an alert for an impending departure from the lane, which requires strict real-time performance. The traditional methods pose high requirements for the real-time inference speed of the model, which cannot be achieved in most environments. The real-time demand is the biggest bottleneck limiting the performance of lane detection in traditional methods, thus leading to the need for a real-time and high-accuracy model.

4.2. Deep Learning Approaches

In, authors present a computational network for lane detection, called REcurrent Shift Aggregator (RESA). This model combines the feature information of spatial and temporal domains to resolve lane ambiguities in challenging urban scenes. In addition to space-variant behavior, the design alleviates visual ambiguities by modeling motion information using stacked gated recurrent units. Empirical analysis on two public datasets demonstrate that the proposed network achieves state-of-the-art performances in different road scenes, i.e., real-time execution and good lane detection even under challenging conditions. On the dataset analyzed in this paper, RESA achieves top performances among other methods, including both state-of-the-art deep learning-based networks and traditional ones. Later, the design of the RESA network was changes to include hierarchical feature processing, scene semantic information and better spatial-temporal context and is implemented in the new architecture by introducing a spatial-temporal aggregator unit, i.e., the STA.

Self-driving cars typically include lane detection and road segmentation modules [7]. While traditional techniques have been used widely, deep learning methods such as CNN, EL-GAN, and VPGNet have increasingly revolutionized road-segmentation applications in urban environments [18]. Increased computational resources and advancements in deep learning methodologies has enabled researchers to build lane-detection and road-segmentation algorithms, which can efficiently segment lane markers, as well as curb and road divisions in urban environments [3].

5. Road Segmentation Algorithms

The Res-Road Network has an attention module responsible for obtaining the global context of the scene, which will largely promote the detection performance. Though deep methods are superior to traditional methods in terms of accuracy and gradually shorten the running time, deep models still face various challenges, especially lack of generalization, in the road segmentation scenario. Summarized existing methods for road segmentation. For example, some network architectures rely on a large number of labeled samples, but there is a problem that the amount of annotations required is large and that it is difficult to cover all environmental conditions. Therefore, some scholars proposed a new instance segmentation algorithm that is end-to-end trainable, named LaneSeg. It aims at simultaneous multi-lane detection and dense edge pixel labeling method for road segmentation [11]. It includes two submodules (edge subnet and instance segmentation subnet) for lane edge and instance segmentation subtasks to preserve the lane boundaries and detailed instances seamlessly, where the edge subnet is an adjacent subnetwork to the instance segmentation subnet.

Road segmentation remains one of the main and challenging research directions in autonomous driving. In this section, we outline new road segmentation technologies, starting from the concept of road detection to segmentation. Segmentation technologies have been broadly classified into two categories: traditional methods and deep learning methods [2]. The sensor-fusion method has been studied extensively in traditional methods, which fuses some features of images from laser radar (LiDAR), camera, and 3D data. These features provide primitive texture and geometry information of the road that are very beneficial for detecting results. Particularly, it included a method fusing LiDAR points using color image information. Fab network consists of four components: two sets of feature embeddings and two one-dimensional convolutional neural networks [6]. Each of them outputs a series of binary masks, which stand for ego-lane and several surrounding road regions simultaneously, where distanceor motion cues are incorporated into the model to gain a consistent understanding of the road.

5.1. Semantic Segmentation

In this section an analysis of the architectures for lane detection and be made by defining a semantic segmentation network capable of identifying the road in all the encounters of the KITTI raw dataset. Three CNN models, used as encoders, object of this section, AlexNet, VGG-16 and UNet 450 will be analyzed able to segment the road in images showing both straight road and involving turns. For each CNN analyzed will be shown the results of the semantic segmentation. Finally, we end this section by comparing the performance of all three models. These promising results allows an initial evaluation of the effectiveness of vertically-proposed models and their capacity to generalize [19].

Despite the great progress in unsupervised and supervised depth estimation methods, depth estimation in critical scenarios remains a challenging unsolved problem. Different failure scenarios from stereo methods are depicted: (i) catastrophic failure due to images that contain pure recurring color patterns or uniform regions and (ii) meaningful displacement (occlusions). To address these drawbacks a novel algorithm is proposed within the probabilistic framework of Markov random fields (MRFs) that combines intensity and gradient information into a single energy functional. Such functional is enforced globally with smoothness constraints, enforced locally to enable sharp intensity transitions and used to enforce the data term. In the experiments, the proposed method behaviour is evaluated in different scenarios achieving accurate depth estimation while over-coming important stereo methods failure cases for both benchmark and critical image pairs [20]. Neural network architectures have demonstrated to be an effective solution for identifying the position and the type of lane markings. In previous works, a Data-Driven approach (D3) was adopted that is trained on synthetic data and tested on real-world images. This method was based on the V-NET architecture for lane marking recognition through an end-to-end Semantic Segmentation CNN in a multi-dataset model. It correctly retrieves the position of the lane markings in space, without applying laws and metric properties, and it is able to generalize in a multi-sensor and multi-mission context [15].

5.2. Instance Segmentation

Several lane segmentation strategies have been presented for the driveable-road region, while this instance-wise labeling subtask has recently gained recognition in the field of deep learning for autonomous driving [5]. We can briefly summarize existing instance segmentation efforts by mainly referring to how instance level operational units are combined to construct neural networks. Besides, fine-grained positional orders like horizontal-top, horizontal-middle, and horizontal-bottom, can also be directly derived to build an offset-aware image encoder. Most existing instance level outputs from these combinations technologies in the early stage just group individual lane components together, resulting in an inexact final instance segmentation. In fact, each lane category with its distinct numerical preference will cluster label with its own position-level unit and emissions. For enough lane-count-specific sub-stage settings, a step-wise relative likelihood of pixels along the ordering direction of this specific lane category can also be inferred.

Instance segmentation, also called object-level semantic segmentation, provides semantic labels along with instance-level identification for the objects in an image. This subtask has recently attracted attention in the lane-detection and road-segmentation tasks to deal with multiple lane lines [7]. Early works on image-wise labeling simply consider each color or brightness change on the road surface as a specific lane segmentation class instance, and they employ different shallow or deep convolutional network settings to recognize lane boundaries and predict the corresponding confidence map in an image-level manner. Since early methods based on instance proposals require a huge amount of memory and inference time, center-based methods progressively predicting instances on the dataset first propose representative center definitions to select target pixels for further instance recognition. Variants of this method achieve a trade-off on accuracy and inference speed. Modern neural-network-based methods resort to end-to-end structure or heavy post-processing, which both call for new solution designs to alleviate the overhead.

6. Datasets for Training and Evaluation

The following research community employs two main automotive datasets for methods comparison for lane detection: the U.S. CULane dataset, comprising only urban roads such as urban, inner-city, local, ring roads, and the Chinese TuSimple dataset, where the database is diversified into traffic complexity, so the method comparison is comprehensive and multifaceted. For automation, there is a specific Activities of Daily Living (ADL) working database for lane detection. However, the ADLi dataset is still not sufficiently suitable since it does not take into account dividers, other vehicles and left-side lanes but it only contains right-side lanes most of the time. So that, our method can be also used as a real-time application if the training base is changed.

Lane detection remains a critical automotive function in the context of autonomous driving, and especially in highways where frequent lane changes and merges occur. Researchers have investigated several lane detection algorithms: traditional methods that use manually designed features [21]; and more recently, methods using deep learning, which automatically extract relevant features from raw data through artificial neural networks [6]. Two databases are commonly used in the automotive community: the CULane and TuSimple databases, focused on urban driving. This work operates on the CULane database, while extending in the highway context with the particularity of an explicit 'prior knowledge' of lanes position and road geometry in general [22].

6.1. Commonly Used Datasets

With that in mind, it is not vigorious enough for the computational demands of real-time environment which showcases the cop cars to execute real-timne on micro fringes and additional motion, reflection and brightness changes. The detailed insights we provide in conjunction with our evaluations should be beneficial to those who take similar datasets into the future to conduct the challenging task of lane segmentation. The trained segmentation model can be seamlessly transferred to additional single-frame forward deployed solutions through 3D information from a single UV space with easy implementation. The annotation on the AMU Radar-Velodyne data is performed, considering several criteria for the accurate localization of the points detected. A category and localization is annotated for more than 10k radar points across 5 types and 11 color segments for Velodyne points. Our dataset is original and is extracted from the real place and the CAD simulation origination of the cars in the simulator. It can be easily supplemented and supported to the semi structured, semi autonomous or fully autonomous cars, and trucks [11].

In the context of showing how we analyzed past success stories and removed weak points, we aimed at excluding their major shortcomings. We achieved that by capturing some streets with roads of both types and require minimal preprocessing for managing the missing link problem during the lane detection stage. To further increase the elevator style legibility, we enforce colour heldout in our physical splitting. The most involved nodes, some of which may be modelling background artefacts, may end up being removed. Additionally, the ego vehicle remains relatively stationary which allows eliminating vehicle dynamic representation. With these sensing hardware and previous success stories, one can consider the proposed ANNA dataset as second, more difficult training phase requiring essential actual experiences for the model so as to better generalize for comparison of the high accelerations of the identifier. Another motivation is that I wonder whether the better segmentation of the roadway might not be required for performing sensor fusion when observing the shape of the occupied space around the vehicle. Based on this I wonder how the predicted path distributions by ANNA can derive into a smaller latent space. It would be interesting to train another model better being able to not only perceive a large amount of homogeneously completed intersections but also substantial amount of high accelerations with different number of lanes and driving conditions [23].

6.2. Data Augmentation Techniques

A technique that would greatly increase the quality and diversity of the dataset would be to generate new images using world simulation. However, this technique is not able to generate enough diversity in the training datasets. Thus, to simulate the changes in weather and time of day, the authors of added in the input of the network the normalized RGB channels of the image and its corresponding colored birds' view. This way, the network is able to learn to ignore color and lighting information when detecting lane borders and focuses on shape information. They still use classical data augmentation techniques such as rotation, scaling, flipping and adding noise with a very low probability. To simulate lane absence, occlusions and others, the authors retrained the networks on a dataset with much more diversity. This dataset is composed of new images using new kinds of augmentations on the clean dataset, which added cases such as, missing information, sun occlusion, large cracks, small cracks, road debris and rain occlusion. This way, the network is able to learn new representations, such as horizontal patterns corresponding to rain on the camera lens, as shown in.

The objective of data augmentation is to generate new inputs from the training data without changing the expected outputs, to increase diversity and to improve the generalization capabilities of the model. Some common techniques include flipping, rotating, skewing, adding noise, scaling, cropping and changing the brightness of the images. With respect to lane detection and road segmentation tasks, the main challenge is to generate new examples that account for the change in perspective, lighting and weather. It is also known that the performances of the models depend directly on the diversity and quality of the training dataset [1].

7. Performance Metrics

Recently, deep learning approaches have been used for both lanes and road segmentation, since deep learning is known to provide a method that is superior to old-fashioned, handcrafted rule-based methods [5]. In this section, we introduce several common quantitative metrics for evaluating models used for lane detection and road segmentation. The real-time performance of these lane detection and road segmentation methods is directly related to their integration in self-driving cars. Therefore, we also discuss the important real-time performance metric that is essential for model production [4].

Lane detection and road segmentation in self-driving cars are Fundamental Operations for ensuring the safety of vehicles. They are also critical for making informed decisions by the self-driving cars concerning driving behavior and path planning. Thus, these tasks necessitate high detection accuracy, especially at nighttime and during rain, fog, or snow events [12]. Due to the ever-changing road conditions, lane detection and road segmentation need to be highly responsive, accurate, and robust to noise, which places a demand on real-time performance.

7.1. Intersection over Union (IoU)

For lane line and road segmentation, the denseness and ordering distinction are important properties that IoU automatically highlights. The IoU value will be 0 when a lane line is not predicted. When the lane line is predicted but its center deviates from the ground truth center, IoU > 0 because big overlap areas with low ordering values still exist. Moreover, considering the dense property of lane labels, if the detector fails to detect a fragment of the lane line, the corresponding predicted fragment will be more ordered, which may also affect accuracy. Thus, the optimisation of IoU takes into account lane pixels and their ordering relations. The deep structure of a neural network can consequently inherit IoU optimisation, so the IoU becomes regularisation for learning in deep neural networks. We introduce the IoU-aware Loss (IoU-Loss) to realise the network using IoU as our supervising signal. The IoU minimum maintains high quality detection and corrects the predictor when regression errors occur. We build IoU from feature levels to overcome the defects due to gradient disappearance. The IoU-aware Loss connects the advantages of both focal loss and Lovász-softmax loss [24].

The intersection over union (IoU) is a metric for determining well the predicted lane lines overlap with the ground truth. If P and G represent the predicted and ground truth sets of lane pixels for a specific category, the IoU is defined by their intersection divided by their union. IoU is widely used in the field of instance segmentation and semantic segmentation to robustly evaluate detection performance. Moreover, the mean IoU (mIoU) calculates the average IoU over all categories and is used as a unified metric to measure segmentation precision [5].

7.2. Mean Intersection over Union (mIoU)

$tIoU=(DnumBoss+DasDashedLnesDashed)(DnumCamPoints+Das\square edCamPoints+DashedCamPoints+DnumCamPoints)$, where the TP, FP, and FN are connected with the following notations: Dn and Dd are true positive and false positive, respectively, while Cn denotes true negative [5].

$IoU=(TP)/(TP+FP+FN)$, where TP, FP, and FN stand for true positive, false positive, and false negative, respectively. Hence, the recall of the lane marks is expressed by the IoU, as tIoU:

Ref: 5d992db3-ffb5-4aaa-9a9b-34d401d3cf9d The second part of this evaluation metric consists of computing the mean Intersection over Union (mIoU). The mIoU simply calculates the average IoU over all classes. It can also be expressed as the average of the true positive rate (TPR) or recall, and vice versa. Our method extends mIoU to lane segmentation, predicting monochromatic (one-channel) segmentation maps, where the background class refers to the non-lane region [25]. The lane marks classes consist of the lane boundary marks, the dashed lane line marks, and the solid lane line marks [26]. Our mean IoU normally falls within the ±10% range. An IoU measure reads as follows:

8. Case Studies

In the second case study, the given problem was solely the lane detection task and the aim was to identify the existing lanes including missing parts. This problem arises in both urban and highway environments and is another typical environmental perception task for autonomous vehicles. The end-to-end lane detection algorithm utilized the master architecture, Pinet, which is successful achieving high lane detection performance on urban and highway environments. All input data are processed with this Pinet module, and then we extract lane center directions, and estimate lane types. Each lane has been extracted completely by the network and outputs lane width and heading angle. The code for this end-to-end lane detection scenario has not been published yet. [7]

In the first case study, the tasks of detecting lanes as well as segmenting roads in a 3D point cloud of an automotive LiDAR for the Middlebury dataset were solved by this type of architecture. The LiDAR data was used directly in the network training taking advantage of LiDAR data advantages and not being limited by the disadvantage of LiDAR data unlike existing CNN-based methods. This is a SN-based end-to-end processing of the continuous point cloud segmentation for lane detection and road segmentation. EL-GAN framework has been used for lane detection and PROBE for road segmentation. The overall GMIOU is 66.1% for the 3D road segmentation with the 27 classes and 64.6% for 3D lane point clustering with the three classes. precision and recall of 68.5% and 90.9% for the road as well as 48.4% and 50.1% for the lane detection.

After discussing the important issues of lane detection and road segmentation, we present two case studies in this chapter. One of the case studies covers the lane detection and road segmentation of the 3D point cloud using an end-to-end neural network. The other case study investigates the performance of a CNN-based lane detection method for the road detection task. The presented network architectures and the lane detection and road segmentation results were previously published in the following articles, collected in this work of thesis for the sake of di erent readers with diverse backgrounds in data or sensor fusion: [4].

8.1. Review of Existing Architectures

End-to-end lane detection networks with minimalistic architectures for limited computational budget have been receiving considerable attention in the industry because of high accuracy obtained using deep learning-based lane detection systems. DeepLabv3plus, a powerful deep learning architecture for segmentation was carefully analyzed for its architectural characteristics and its utility for lane detection and drivable area segmentation task in autonomous vehicles [27]. As an important contribution, a new segmentation architecture, ESPNetv2 which is considered as successor to ESPNet, was proposed with new module called attention gate decoder module. The ESPNetv2 is a twinnet-based solution with intermediate supervisions, much faster than convolutional neural network (CNN)-based solutions and better than state-of-the-arts ESPNet and DeepLabv3plus for lane detection and drivable area segmentation task. Making the ESPNetv2 further efficient and lightweight for the lane detection, the segmentation decoder network was transformed into an appropriate structure for road scene understanding to showcase an alternative method for lane detection.

Although deep learning has been influential in visual perception, algorithms for efficient end-to-end learning of complex maneuver tasks still require significantly larger data sets than traditional algorithms. The development and performance of deep learning-based solutions for vision tasks using data fusion pipelines during acceleration, such as steerable filters with applications to lane detection, were investigated. A LiDAR filter, MAX2131, received data from the forward-facing lidar and reverse-engineered to produce the Adamski filter response in the steering space, while a deep neural network using lane departures was deployed at the output layer [2]. The good experimental results obtained with the AV dataset confirmed the correctness of the implementation, with an average runtime prediction of 60.18 Hz, but the transferability to new data required more study. Thus, new data sets were proposed.

8.2. Comparison of Performance

Current content-aware data augmentation and/or synthesized training data through generative adversarial networks (GANs) techniques have also been developed for annotated data set strengthening. As it can be understood from the experimental results, the proposed YOLO (You Only Look Once) lane detection method with modified deep learning-based CNN architectures and training layer sequence have been obtained better than the original YOLO v3 lane detection method on the shared lane detection database. Although the YOLO v4 lane detection achieves a success of 99.9% in the training process, it has a delay-sensitive convolutional neural network architecture and speed-based crack neural network representations that can prevent false negatives generation on CNN-based strong lane structures. However, since most road and lane structures exhibit DRNN-based non-linear representation, the selected lane structures have been questioned if the advanced RNN architecture (e.g., Bidirectional-RNN) gives additional information. [10] Reinforcement playing method has been used for refining the architecture layers in the line detection pipeline, which can maximize performance rate and minimize the training accuracy between convolutional and rectified linear layer sequence.

According to [13] as it can be observed from the experimental results, the line detection via deep learning achieved a success of 98.8% in the accuracy rate and 96.9% in the speed, while the previous work based on the pixel-wise semantic segmentation of the road images for line detection accuracy was approximately 85%. This success emphasizes that preliminary performance in the deep learning-based methods is quite prominent. Therefore, the two important reasons why line detection-based neural network architecture is feasible are also highlighted. The high number of tunable parameters in the already trained neural networks that are used for the object detection and semantic segmentation helps to draw accurate lines and increase the detection rate. In contrast to the training of traditional and object-based neural networks from scratch on the significantly small or medium-sized data set for relatively simpler binary decisions like obstacle detection, tracking, vehicle detection, lane detection, and so on, using line parametric representation in the line detection provides a high success rate for the image's and/or video's domain transformation.

9. Challenges and Future Directions

Since pedestrians are usually on the roads, we will need to cross the footbridge to avoid pedestrian detection, which leads to a large range of pedestrian misalignment about the footbridge. A favorable way to retain the existing lanes in urban roads occurs when we build a pedestrian bridge[lane carpet, grove, mask, trunk] in the feature alignment of the attentional layer, while the correct big region of pedestrian. The pedestrian's location is avoided, making the model focus on object-level tasks. The multi-lane object-level module and the pedestrian-level features are thus independent. For the multi-task performance analysis in a larger-than-IVS-SEG urban environment, the results fully demonstrate the effectiveness of our proposed approach..

One of the major challenges faced during lane detection in an urban environment is the existence of multiple lanes on the road, often with fades or small "width" lanes[e.g., 1]. Due to the connectivity of lanes and the limited resolution of the road, the multi-lane detection task is more challenging than the single-lane detection. Moreover, in the lane detection dataset, the properties of a pixel include not only the ownership of the lane but also the position in lane width, as shown in Figure 4. To sum up, in this paper, we empirically find that metaphytes detection promotes multi-lane detection. The first step detects a metaphytes and classifies it as a class, such as "left-turn-left, right-turn-left, right-turn-right, left-turn-right", etc. , where the left and right turns represent the direction. These classes are identified first before locating the exact positions of the individual lane components..

The field of autonomous vehicles has evolved significantly over the years. Yet, there are some gaps in lane detection and road segmentation that need to be worked upon. Here, we highlight some of these challenges and discuss potential solutions to resolve them more efficiently.

9.1. Robustness and Generalization

The novelty of our work is to design scale-specific and robust detection and flatness strategies to solve environmental changes and diversity problems for lane detection. In response to the generalization problems mentioned, for the augmentation strategy, the attributes of each component, which constitute the proposed components, reflect the different environmental variations. Taking into account the challenges laid out above, powerfully utilizing and jointly optimizing our method with other features would provide more holistic understanding of these tasks and potentially result in enhancement of current benchmarks. Nicely incorporating the powerfulness of the core neural networks would provide further evidence. Considerable improvement is reported on the virtual KITTI dataset for road segmentation and 80% of the TuSimple dataset for the lane detection task (ref: 1a665d3f-ae88-4287-9f6b-87fabb59fa0f).

A lane detection model is required to predict lane markings accurately. The major challenges include the influence of weather and illumination changes on the model robustness as well as the generalization of a model solution to large-scale variations in new scenes. (EL-GAN) provides a lane mask policy to produce a structure-preserving output in a top-down manner, thereby improving the performance of the model (ref: 425a66b2-a389-4fb4-9877-d13b4d712442). Although this topic has been researched from various perspectives, such as the application of end-to-end methods or the combination of CNN+RANSAC, it is essential to develop an undiscovered method that overcomes weaknesses in previous methods. In previous studies, end-to-end or regression based network architectures are used as popular methods for lane detection. These methods have several common weaknesses such as ignoring structured lane patterns or ignoring flexibility in model structures to solve tasks with different characteristics (ref: 55b165bc-4d99-4801-a321-cf3820a6b84b).

9.2. Real-Time Processing

Lane detection in autonomous driving area is an important task and its accuracy and speed is crucial in terms of many safety-critical applications like self-driving vehicles, advanced driver assistance systems(ADAS), traffic control, car-following, collision detection systems of the vehicles and so on. LaneParallizer is the first trainable detection method that can detect lane-parallel vehicles and is robust against occlusions, visual background clutter, and lane occluders. LaneParallizer outperforms prior vehicle and lane detection approaches by a significant margin and surpasses the state-of-the-art occlusion handling strategies in terms of precision and recall. The LaneParallizer is robust to arbitrary occlusions in city scenes without using occlusion annotations or manual feature engineering. The pipeline has been designed to reduce redundancy in depth maps, decouples vehicle segmentation from depth-based feature extraction, and features efficient processing of unoccluded free space.

Lane marking detection in autonomous driving is a significant research topic due to its simplicity, low cost and accuracy [28]. Convolutional Neural Networks (CNN) have been widely used for this task recently [21]. In this work, the deep learning-based lane detection method introduced in, we forecast future lane lines by designing and training a U-Net based CNN on the DiscoBox dataset [6]. The system is also optimized for fast real-time processing using FPGA acceleration which is available easily in many marketing road vehicles. Unlike existing systems, the system proposed in this paper considers lane markings not only at the same frame but also over multiple frames, enabling it to forecast future lane lines which is crucial for autonomous driving in fast moving vehicles.

10. Conclusion

This researcH includeD an analysis of the tesT results relaTeD to The use of bottlenecks (masks for The filter outputs for The textures noT reLevanT for The networK), an atTempT to miniMize These losses and work only with The primary textures. The bottlenecKs turn out to be Ideal for simplifying The complexThe sen sory input anD making The netWork much easier To tune [14]. On the other hanD, we have shown THaT the qualiTy of The ouTpuTs gener-aTeD by The netWork can be improved by The inclusion in The nT dimensions of The TwoD standards, which descriBe The grayscales of The colo rs appearing close To each oTher, of The ouTpuTs of The lowOposiTe scale equivarianT operator anD of The g" channel of each Moire. This alloweD us To operate in space, and noT in frequency.

Lane detecTion is a key componenT of road segmenTaTion in seLf-driving cars and advanced driver assisTance sysTems (ADAS) [5]. We have presenTed a sysTem basEd on a redesigned version of a popular neuRal network for lane detecTion, With an improveD data alignment procedure. The key idea of oUr method is to utIlize feaTure map availability and penetrative vetoes to selecteD represenTations at differenT scales. Our moDel achieves increaseD performance anD a faster execution time [3].

10.1. Summary of Key Findings

U-Net, or more generally FCNs (Shelhamer et al., 2017), have built up leading studies in road detection, with their capacity in edge retention on complex objects and scalekeeping that can initially assess boxes of arbitrary sizes in the image. Because the unskipped paths in the U-Net path prompt shoulders, redundant curves were employed to better improve road detection. The shapes and locations of the ground truth lanelines are more discriminative than the scenes that may or may not have lanelines, which allows for better guidance for complex learning tasks. Unfortunately, non-attention mechanisms that focus purely on road segmentations are affected by lane appearance noise and distracting elements that result in target information loss (Turi et al., 2021). A weak pixel and high order information decoder were incorporated into U-Net and spatial attention WrapNet+ and EffNet improved, with the nets significantly achieving the highest Focal-Loss-plus-Dice-Loss segmentation accuracies for GTD4-1 and GTSD4, respectively, in their data groups.

Semantic segmentation in self-driving tasks involves pixel-level classification to delineate objects and their boundaries. In the case of the drivable area and lane segmentation for pixels of the road IRIS, DL methods such as TwinLite parsed all of the following types of lanes: dotted, crossed, solid, and lane edge locations, achieving a maximum of 54.1% accuracy for multiple lane detection [11]. While semantic segmentation of lanelines has recently come to the fore with the adoption of parallel architectures such as UNet (Ronneberger et al., 2015) and ResUNet (Alom et al., 2020), real-time robustness requirements have not yet been fully met for modern AI trends. A strong recarving network, TwinLiteNetPlus, combined with ResNet-34 and ResNet-101 backbones, was thus purpose-built [12]. With only image resolution being altered for the ICLane and Tusimple datasets, the experimental comparisons demonstrated meaningful improvement for the intended object goal that benefits from the resultant positive effects, with the perception that the road area is also smoothed while ensuring 1.6 and 1.5 turned ups (up to 11.6 ms) on 1600 × 1200 and 1280 × 720 image resolutions, respectively.

10.2. Implications for Self-Driving Technology

Latest deep neural models allow for exploiting the complete scene information embedded in a single color image; instead outdoor environments demand for several sensorial streams to be fused together to provide a robust perception. A heavily reduced model has been shown to obtain a globally comparable performance (0.998 versus 0.999 AUC score on the commonly adopted CULane benchmark) with a great drop of the number of parameters (-87.7%) and a shorter execution speed up to 10 times faster (49 ms instead of 499 ms). Summarizing, ants can help in intake monocular estimation fusion almost all the scene information, are a robust perceptual aid especially in harsh environments: 1) where lanes are heavily occluded by static obstacles (vehicles, road signs, bridges and tunnels); 2) where lane markers are unrecognizable due to low-visibility conditions (fallen barriers with missing stripes, non-reflective lane markers wet from the rain).

Neural network architectures for lane detection and road segmentation already provide significant improvements over previous efforts in terms of robustness [29], accuracy, and execution speed thus creating new opportunities for autonomous driving applications. The final step in performance validation is the deployment of the DL architectures in a real-world, commercial vehicle. Recent advancements in the related field of public and private vehicles for data gathering and vehicle-tracking indexing databases (monolithic or decentralized) will make it possible to massively exploit positioning information and linked navigation systems (LNS) at a global level to annotate lane and roads. This potentially endless labelled ground truth dataset for actively-driven RGB-videos will allow better learning from real-world data and to use it for any kind of task, exploitation in real-time of this kind of information is required not only for lane detection, but in general for all the ADASs related applications.

Reference:

1. Perumalsamy, Jegatheeswari, Bhargav Kumar Konidena, and Bhavani Krothapalli. "AI-Driven Risk Modeling in Life Insurance: Advanced Techniques for Mortality and Longevity Prediction." *Journal of Artificial Intelligence Research and Applications* 3.2 (2023): 392-422.

2. Karamthulla, Musarath Jahan, et al. "From Theory to Practice: Implementing AI Technologies in Project Management." *International Journal for Multidisciplinary Research* 6.2 (2024): 1-11.

3. Jeyaraman, J., Krishnamoorthy, G., Konidena, B. K., & Sistla, S. M. K. (2024). Machine Learning for Demand Forecasting in Manufacturing. *International Journal for Multidisciplinary Research*, *6*(1), 1-115.

4. Karamthulla, Musarath Jahan, et al. "Navigating the Future: AI-Driven Project Management in the Digital Era." *International Journal for Multidisciplinary Research* 6.2 (2024): 1-11.

5. Karamthulla, M. J., Prakash, S., Tadimarri, A., & Tomar, M. (2024). Efficiency Unleashed: Harnessing AI for Agile Project Management. *International Journal For Multidisciplinary Research*, *6*(2), 1-13.

6. Jeyaraman, Jawaharbabu, Jesu Narkarunai Arasu Malaiyappan, and Sai Mani Krishna Sistla. "Advancements in Reinforcement Learning Algorithms for Autonomous Systems." *International Journal of Innovative Science and Research Technology (IJISRT)* 9.3 (2024): 1941-1946.

7. Jangoan, Suhas, Gowrisankar Krishnamoorthy, and Jesu Narkarunai Arasu Malaiyappan. "Predictive Maintenance using Machine Learning in Industrial IoT." *International Journal of Innovative Science and Research Technology (IJISRT)* 9.3 (2024): 1909-1915.

8. Jangoan, Suhas, et al. "Demystifying Explainable AI: Understanding, Transparency, and Trust." *International Journal For Multidisciplinary Research* 6.2 (2024): 1-13.

9. Krishnamoorthy, Gowrisankar, et al. "Enhancing Worker Safety in Manufacturing with IoT and ML." *International Journal For Multidisciplinary Research* 6.1 (2024): 1-11.

10. Perumalsamy, Jegatheeswari, Muthukrishnan Muthusubramanian, and Lavanya Shanmugam. "Machine Learning Applications in Actuarial Product Development: Enhancing Pricing and Risk Assessment." *Journal of Science & Technology* 4.4 (2023): 34-65.