Journal of Advanced Research in Applied Sciences and Engineering Technology 39, Issue 2 (2024) 166-180

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage: https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Fully Convolutional Network Model Applied Attention Mechanism on Kitti Lane Dataset for Lane Detection

Noor Jannah Zakaria[1], Mohd Ibrahim Shapiai[1,*], Rasli Abd Ghani[1], Nurbaiti Wahid[2], Daphne Teck Ching Lai[3]

[1] Malaysia–Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia
[2] College of Engineering, Universiti Teknologi Mara, Dungun, Terengganu, Malaysia
[3] School of Digital Science, Universiti Brunei Darussalam, Negara Brunei Darussalam

ARTICLE INFO

Article history:
Received 5 August 2023
Received in revised form 29 December 2023
Accepted 15 January 2024
Available online 12 February 2024

Keywords:
Fully convolutional network; attention mechanism; Kitti lane dataset; deep learning model; autonomous vehicle

ABSTRACT

This article investigated the attention mechanism implemented by the Fully Convolutional Network (FCN) Model on the Kitti Lane Dataset. Two attention mechanisms were applied in the deep learning model to improve traffic lane detection for autonomous vehicles. The Kitti lane dataset, which was generated in collaboration with Jannik Fritsch and Tobias Kuehl from Honda Research Europe GmbH, was selected for this study. The results demonstrate that the applied attention mechanism can effectively improve the network's feature representation on lane markings. Furthermore, this approach can improve the weighted information of lane line targets while decreasing irrelevant information. As a result, the proposed technique improved, obtaining more than 95% accuracy. Subsequently, the attention mechanism was implemented in the FCN model architecture to enhance the lane-detecting model. As a result, in the future, more comprehensive ideas, such as combining the FCN model with Transfer Learning, will play an essential part in investigating the improvement of lane detection areas.

1. Introduction

Engineers and researchers have dedicated their efforts to enhancing automobiles in multiple aspects since their inception, encompassing safety, handling, performance, efficiency, and durability [1]. One of inventions in the automobiles field is Advanced Driver Assistance System (ADAS) which equipped with several modules. One of the important modules of ADAS in autonomous vehicle is Lane Detection. The lane is a traffic sign that separates a road and ensures that vehicles are driven safely and efficiently. Lane detection is a method for automatically identifying road signs to guarantee that vehicles stay inside their designated lanes and do not collide with other vehicles. It has contributed to autonomous driving. Consequently, accurate lane identification enables the

* *Corresponding author*
E-mail address: md_ibrahim83@utm.my

https://doi.org/10.37934/araset.39.2.166180

autonomous vehicle to make many decisions and judgments regarding its location and state, thereby ensuring safe driving [2]. However, using lane detection algorithms is difficult due to the enormous diversity of lane markings, the complexity of the road conditions and environment, and the intrinsic narrowness of the lane [3]. Nevertheless, research has led to the development of dependable lane detection systems [4].

Different hand-crafted methods, including geometric modelling and classical approaches, have been applied to detect lane markings. However, most traditional detection algorithms use pipelines that consist of pre-processing the images, extracting lane features, fitting the lane model, and tracking the lane line [5]. First, image pre-processing is used to decrease the amount of noise in an image. The characteristics of lanes are then used in the feature extraction procedure to extract lanes. Following that, the lane model is fitted and tracked using various approaches.

Several techniques have been utilized in the past for feature extraction, including Inverse Perspective Mapping (IPM)/Perspective Transform, filtering technique, image district extraction, morphological operators, neighbourhood searching-based feature points, grayscale, thresholding, clustering, heterogeneous operators, and sliding window [4-13]. These methods help to minimize noise and make it simpler to identify lanes. The next step is commonly fitting the lane model with a line segment detector (LSD) and fitting-based approaches such as the least square methodology, B-splines, quadratics, polynomials, hyperbolics [11,14-18]. After that, the three approaches that are used most commonly in tracking road lane detection are the Kalman filter and the parabola equation [19,20]. In addition, tracking is the post-processing phase utilized to correct for variations in illumination [4]. Consequently, tracking assists with detecting improper occlusions, which might be caused by inadequate lane markers [21]. The traditional methods, on the other hand, include a more challenging and hand-crafted procedure, which results in a significantly more extended period to complete the entire process.

Then, due to the growth of Artificial Intelligence (AI) systems in recent years, the recognition of lane markings has become substantially more accessible, significantly faster, and particularly more effective. In addition, there is no longer a requirement to carry out procedures using hand-crafted methods. AI refers to the practice of computers, most notably computer systems, attempting to mimic the cognitive processes of humans [22]. Most of the AI strategies utilized in lane detection may be classified into two broad categories: machine learning (ML) and deep learning (DL). Because of its superior performance in either classification or detection, the DL technique, which uses image frames as input to the network algorithm, has surpassed ML as the method that is considered to be the most widely used [23]. This is the crucial factor contributing to the DL method's remarkable increase in popularity.

In addition, some scholars advocated applying the DL method as a stand-alone strategy [24–27]. In contrast, the others advocated combining this method with one or more other techniques. The integration of this network serves as its primary objective, the improvement of the effectiveness of the network in demanding situations concerning the detection of the lane mark. The combination of DL with geometric modelling is one example of the integration of DL with another method [13,28-31]. Another example is the integration of DL with ML, also integration of DL with DL which is known as the two serial DLs [32-38]. Aside from that, in more recent times, an innovative approach for integrating this method has been proposed, and it involves the combination of DL with an attention mechanism [3,4,39,40]. However, this is the most current state-of-the-art method that has been suggested, and there is still an opportunity for additional research.

In recent years, attention mechanisms in computer vision have been a significant area of study due to their capacity to improve model performance by a substantial margin with low parametric overhead. The most significant interpretable aspect of any attention mechanism is that it allows the model to emphasize more essential and deterministic features [41]. Nevertheless, attention methods increase parametric complexity like any trainable layer or module. This work aims to demonstrate a novel technique for applying attention to computer vision models at the cheapest parametric cost while performing well in competitive tasks such as lane detection. This process has become an intrinsic component of modern deep neural networks to boost the generalization capabilities of these architectures. Many unique attention processes have been proposed in computer vision over time. Squeeze-and-Excitation Networks, Convolutional Block Attention Modules (CBAM), Bottleneck Attention Module (BAM), Global Context (GC) Nets, Attention Augmented Networks (AANets), and A2 Nets are a few examples [42]. However, most of these attention methods involving any type of channel have substantial drawbacks, including a significant increase in parametric complexity, a reduction in dimensionality, and the absence of cross-channel interaction (CCI).

This research proposes the lane detection method using Fully Convolutional Network (FCN), incorporating an attention mechanism architecture to address these limitations. Consequently, the DL approach has good generalization potential and learns the driving environment’s essential features. However, only a few works of literature have investigated lane detection using the attention mechanism with deep learning, mainly based on a fully connected layer. Furthermore, the convolutional layer in FCN will further emphasize salient features derive from the attention module. Therefore, the attention mechanism emphasizes the significance of network learning features.

The rest of the article is as follows: Section 1 discusses previous methodologies relevant to this research. Then, section 2 elaborates on the proposed method, while Section 3 briefly discusses the outcomes of the proposed experiments and modeling on the Kitti lane dataset. Finally, section 4 concludes with a summary of the study's findings and recommendations for future research.

2. Methodology

2.1 Kitti Lane Dataset

A dataset with open access and standard evaluation means has been made available via the Kitti lane dataset. In the field of autonomous driving, various datasets have been published online, including those from TuSimple, Caltech, and Cityscapes, as well as the Kitti dataset for the lane dataset [43-46]. However, when it comes to all of these, the Kitti lane dataset is one of the most widely used ones to assess the system's performance in autonomous vehicles. It includes information about the road scenes, such as laser points, colour images, and stereo views [46]. Experimental data, specifically the Kitti dataset, are utilized by many researchers to evaluate the effectiveness of the proposed method utilized in lane detection. This study's proposed lane detection was carried out using the Kitti lane dataset for only colour images from the Kitti lane dataset. With the use of this dataset, evaluations and comparisons have been carried out. This dataset was developed in collaboration with Jannik Fritsch and Tobias Kuehl at Honda Research Europe GmbH [46].

The datasets include three distinct road scene categories, which are referred to as urban unmarked (UU), urban marked (UM), and urban multiple marked lanes (UMM) [46]. This road and lane estimation benchmark include 579 images with a resolution of 375 by 1242 pixels. Two hundred eighty-nine training images come with their corresponding ground truth. The remaining 290 images are test images depicting various road scenes, such as urban, rural, and highway settings. The images have been manually annotated to create the three types of road scenes over ground truth. It consists of two different terrains: road – the road area, also known as the structure of all lanes, and lane – the ego-lane, also referred to as the lane the vehicle is actively traveling on (only accessible for category "UM"). Both are available for the road area, also known as the composition of all lanes. In addition to that, the annotation is only presented for the training dataset. A total of 600 annotated training and test images covering a diverse range of urban road sceneries are included in the entire dataset. Next, this benchmark provides a ranking of methods based on established metrics, including MaxF, which is the maximum F1 measure, AP, which is the average precision, PRE, which is the precision, REC, which is the recall, FPR, which is the false positive rate, FNR, which is the false negative rate, and F1, which is the F1 score. The whole collection of data was analyzed and then sorted into the three groups demonstrated in Table 1.

Table 1
The Kitti Lane Dataset includes the scene category and the total number of training and testing datasets [46]

Scene category	Training	Testing
UU (Urban Unmarked)	98	100
UM (Urban Marked two-way)	95	96
UMM (Urban Marked Multilane)	96	94
Urban (all)	289	290

In addition, the Kitti lane dataset includes synchronized images, annotations images for only training data that are carefully annotated, and the calibration parameters that are relevant in the process of providing a comparison framework with various techniques. In addition, this public database includes information on road scenes that was gathered using a variety of sensors, including an INS (OXTS RT 3003), a LIDAR (Velodyne HDL 64E), two color and two monochrome cameras with 1.4 megapixels each (Point Grey Flea 2), and four varifocal lenses with a focal range of 4-8 millimeters each (Edmund Optics NT59-917) [46]. This dataset also includes a web interface that may be used to analyze several road recognition methods in the 2D Bird's Eye View (BEV) space. As a result, the Kitti lane dataset served as the basis for the evaluation in this study of the model of FCN that utilized both mechanisms of attention. The instances of the Kitti lane dataset and associated ground truth used in this investigation are displayed in Figure 1.

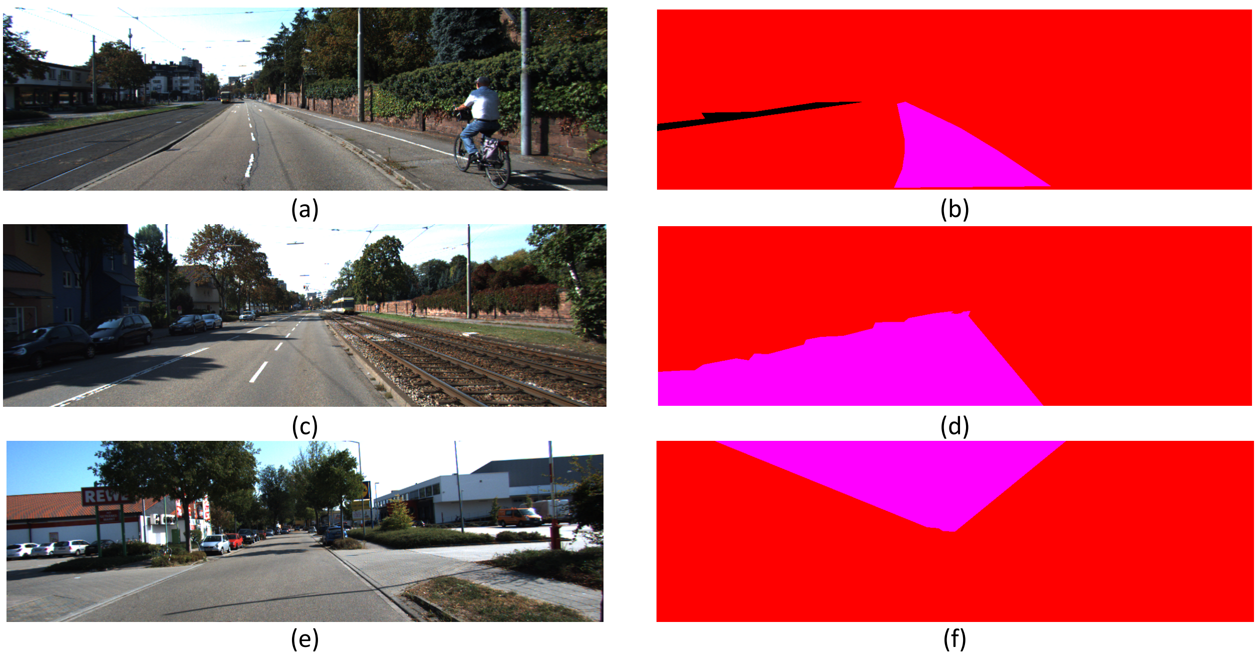

Fig. 1. Exemplifications of the Kitti lane dataset implemented in various metropolitan settings (a) Kitti lane UM categories dataset (b) UM training image ground truth (c) UMM Kitti lane dataset (d) UMM Kitti lane ground truth (e) Kitti lane dataset UU training image (f) Kitti lane dataset UU ground truth

2.2 Fully Convolutional Network Model with Application of Attention Mechanism

Details of the suggested framework for lane detection in an autonomous vehicle employing a Fully Convolutional Network with an attention mechanism are presented in this section, as shown in Figure 2. The framework is divided into two models: FCN with channel attention mechanism and FCN with efficient channel attention. In general, the attention mechanism is introduced to the FCN layers to improve feature localization in the feature map, feature extraction capabilities, and network feature representation without adding any reference time [39]. Furthermore, acquiring a sizeable receptive field efficiently captures the information of the lane. In a typical FCN, the feature map is gradually down-sampled for this phase. A novel attention mechanism based on FCN was proposed in this paper to detect road lane markings for an autonomous driving system. The attention gradually improves the ability of the neural network's layers to represent features. Following that, it enhances feature localization in the feature map. Therefore, the attention processes in feature maps can highlight crucial spatial information [40]. Also, the attention mechanism can boost the weighted information of lane line objectives while reducing unnecessary data [3]. As a result, it makes network learning more interesting.

2.3 Channel Attention and Efficient Channel Attention

One of the first works to introduce a novel channel attention method was presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2018 [42]. It is widely regarded as one of the most prominent studies in attention mechanisms. This basic but practical add-on module can be appended to any baseline network to boost performance while incurring the minimal computational expense. There are two primary components in the FCN-based method: the input tensor and the trainable convolutional filters, which contain the weights for that layer [47]. In detail, the tensor is typically represented by a dimensionality (B, C, H, W) that includes batch size, channels, and spatial dimensions [47]. The alphabet of B denotes the batch size, C represents channels, and $H\ and\ W$ are feature map dimensions in which H indicates the height and W indicates the width. Furthermore, convolutional filters extract input data into channels.

On the other hand, different channels may have different levels of representative weight in their data [42]. Before information is passed on to the subsequent layer, it is necessary to assign a weight to each channel based on its significance. This is because some channels may be more important than others. Generating feature maps from the learned weights contained within convolutional filters is the filters' responsibility. While some filters learn edges, others learn textures, and collectively, they learn a variety of feature representations of the target class information inherent in the image of the input tensor [48]. As a result, the number of channels reflects the number of convolutional filters used to discover the various feature maps of the input data.

Moreover, the significance of these feature maps varies. This indicates that different feature maps have varying degrees of relevance. This lends credence to the idea that some feature maps are more significant than others. For learning, for instance, a feature map that only contains information on background texture transitions may be less valuable than a feature map that only contains information on edges. Consequently, on a fundamental level, one would want to produce "more important" feature maps that have greater importance than the feature maps that they are compared to. This lays the groundwork for channel focus. This work aims to emphasize this "attention" on the most crucial channels to give specific channels greater weight than others. The most straightforward approach is to scale the channels with greater importance by a more excellent value.

In the FCN network model architecture, the suggested module employs channel attention and an efficient channel attention block. Because of the diversity of image information in the feature map of the lane, several unnecessary pieces of information should be eliminated before weight computation while still maintaining key texture features to optimize the feature sophistication [49]. The channel attention mechanism represents and evaluates the relevance of each channel using a scalar. For example, let $X \in \mathbb{R}^{C \times H \times W}$ is the image feature tensor in networks, C is the number of channels, H is the feature's height, and W is the width of the feature. Because it must represent the entire channel, only one scalar can be used; the scalar representation in channel attention is a compression problem. The attention mechanism can thus be stated as follows

$$\mathrm{att} = \mathrm{sigmoid}(\mathrm{fc}(\mathrm{compress}\ (\mathrm{X}))) \tag{1}$$

where $\mathrm{att} \in \mathbb{R}^{\mathrm{C}}$ is the vector of attention, $\mathrm{sigmoid}$ is the Sigmoid function, fc is the mapping functions such as fully connected layer and one-dimensional convolution and $\mathrm{compress}: \mathbb{R}^{\mathrm{C} \times \mathrm{H} \times \mathrm{W}} \rightarrow \mathbb{R}^{\mathrm{C}}$ is the compression technique.

Following the acquisition of all C channels' attention, each channel of input X is scaled by the associated attention value

$$\mathrm{Xoutput}_{:,\mathrm{i},:,:} = \mathrm{att}_{\mathrm{i}}\mathrm{X}_{:,\mathrm{i},:,:,} \quad \mathrm{s.t.\,i} \in \{0,1, \dots, \mathrm{c}-1\}, \tag{2}$$

where $\mathrm{Xoutput}$ is the output of the attention mechanism, $\mathrm{att}_{\mathrm{i}}$ is the i-th attention vector element, and $\mathrm{X}_{:,\mathrm{i},:,:,}$ is the i-th input channel.

Next, efficient channel attention is a lightweight, easy-to-plug-in module that can be used in all types of deep convolutional neural networks that efficiently computes channel attention at each point of network architecture introduction. In addition, it is a very lightweight, efficient channel attention module based on enforcing cross-channel interaction without dimension reduction during computational channel attention. Consequently, the parametric overhead provided by efficient channel attention is negligible and directly equivalent to the kernel size k employed in the 1-D convolution layer, resulting in fundamentally inefficient channel attention. Finally, to improve the lane-detecting problem in autonomous driving, it is required to design systems that better use additional data. Consequently, this study proposes employing Channel Attention and Efficient Channel Attention Networks, which improve lane detection performance and make the network function more interpretable.

2.4 FCN Applied Channel Attention and Efficient Channel Attention

The Fully Convolutional Network model was fitted for the Kitti lane dataset, which is used for the learning process. The FCN model uses multiple layered neural networks, each of which, in turn, makes use of mathematical properties to decrease the error of the predicted value compared to the actual value. This enables the model to effectively learn as it is being trained using the data images, which allows it to converge toward a final model. The FCN comprises networks that perform convolution (on the left) and deconvolution (on the right). This network's configuration included the setting of several crucial parameters detailed in Table 2, including the number of epochs, the batch size, the number of classes, and the image shape. In addition, the activation function known as Rectified Linear Units (RELU) was utilized in this particular form of the network.

Table 2
Setting of parameters

Parameter	Values
Number of Epochs	100
Batch size	96
Strides	94
Padding	290
Number of Classes	2
Image Shape	80 × 160 × 3
Activation Function	RELU
Kernel filter	3×3
Max pooling	2×2
Dropout	0.2
Optimizer	'Adam'

FCN consists of multiple layers, including one layer of normalization at the beginning of the network model, seven convolutional layers, three max-pooling layers, ten dropout layers, seven deconvolutional layers, four unpooling layers, and eleven attention layers. This experiment employs two types of attention mechanisms: the channel attention mechanism and the efficient channel attention mechanism. Next, the first network layer consists of the input image with a pixel size determined by height (h), weight (w) colour channels (d). The input images are raw RGB images with 80 × 160 × 3 pixels. The network receives a three-channel image patch as input. The training set and annotations are also patched in the form of arrays. The data is then shuffled to ensure that various images are exhibited instead of merely specific ones. Finally, the labels were separated into lane and non-lane categories. Figure 2 depicts the network architecture, which consists of the following components:

i. The image is filtered in the first layer using a 33-kernel filter with 64 filters and seven strides.
ii. Following this, the normalization process is carried out to improve the accuracy of the network. After it is complete, the convolution layer extracts feature from the image.
iii. After that, at the layer which contains the max-pooling block, the size of the image is halved, and a 2x2 pooling layer is applied; this reduces the amount of memory of the FCN. Finally, after the second, fifth, and seventh convolutional layers, Max pooling is used to decrease the overall size of the feature maps' spatial representation.
iv. Subsequently, a dropout layer is introduced to prevent an overfitting issue. Eventually, the dropout layer value was set to 0.2.
v. The attention layer was then added. The network was initially provided with channel attention, followed by efficient channel attention.
vi. Restoring the image to its original size requires unpooling and deconvolutional layers; consequently, the FCN model was trained to generate a 1x1 convolution of these layers. This was done to facilitate the restoration of the image to its original size. In the end, the 1x1 convolution layer was added in the middle of the network architecture to achieve the following goals: (a) Increase the network's depth and allow it to learn more complex image characteristics; (b)Reduce by a factor of 0.5 the number of channels to accelerate the training process.

vii. Next, the deconvolution layer mirrors the convolution layer or transposed convolution. It is known as a fractional stride convolution, and its purpose is to increase the activation size to match that of the input.

viii. The final deconvolution layer concludes with a single filter because the image only needs a single filter to be returned to the color channel designated as "G." This is the case because the anticipated lanes were drawn in green and then stacked with the red ('R') and blue ('B') channels set to zero to integrate with the original road image.

As depicted in Figure 2, the network propagates through multiple layers, alternating between convolutional, max pooling, dropout, and attention layers. The actual outcome was determined by comparing the input images with the expected output at each layer. The network then down samples by propagating alternately through the deconvolution, dropout, and unpooling layers since it acts as a mirror image. After the model has been created, padding and resizing operations are carried out in order to compensate for the effects of the convolutional and pooling layers. Cross-entropy and the Adam optimizer were utilized as optimization techniques in this instance. The Eq. (3) that defines the algorithm for cross-entropy loss is as follows [50]

$$\mathrm{L} = \frac{1}{\mathrm{N} \times \mathrm{W} \times \mathrm{H}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \sum_{\mathrm{m}=1}^{\mathrm{W}} \sum_{\mathrm{n}=1}^{\mathrm{H}} \log \mathrm{p}_{\mathrm{m,n}}^{\mathrm{i}} \tag{3}$$

where W is the width of the output SoftMax layer, H is the height of the output SoftMax layer, N is the batch size, and p is the probability that the FCN will predict the correct class.

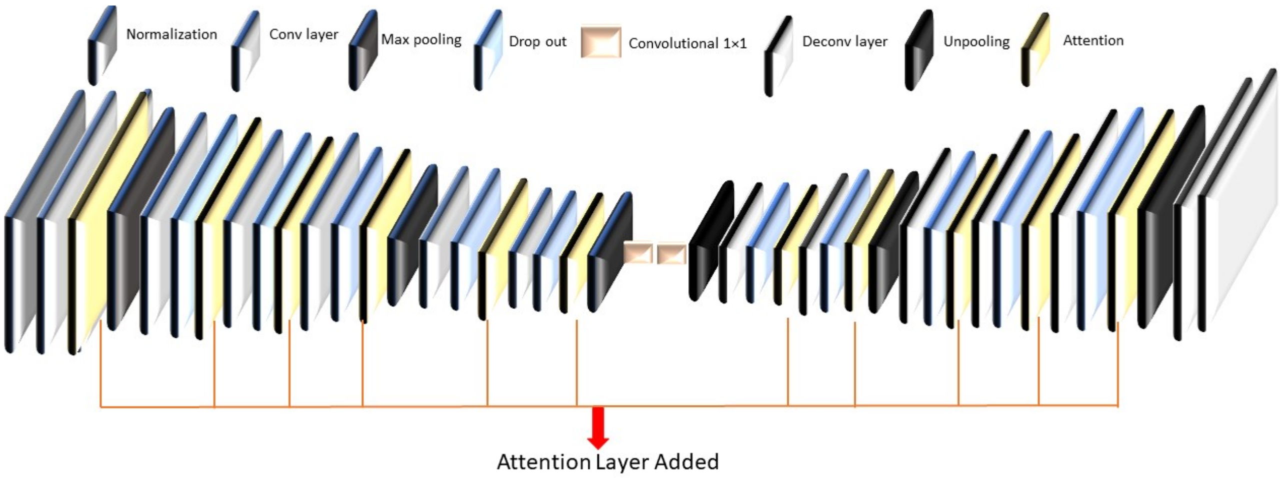

Fig. 2. Model of FCN network architecture for lane detection It includes convolutional and deconvolutional layers, as well as normalization, max-pooling, and dropout layers. In addition, to improve the network architecture's performance, the attention block layer was added after the dropout layer, as indicated by the red colouring

This component is necessary for optimizing the weights and minimizing the deviation between the actual and expected outputs. The weights were modified to meet the optimal requirements of the network. Error reduction will reduce the network's computational time, accelerating the process. The network can handle input images of variable size, with output corresponding to lanes and non-lanes. Next, the network was evaluated based on its accuracy, loss, precision, recall, and F-score during the training and validation phases. Accuracy is the degree to which the network accurately detects the lane. The loss, in the meantime, is the difference between the predicted pixel values of the output lane image and the label for that image. The network used Mean Square Error (MSE) to calculate loss, which is the average of all squared differences. Precision is the proportion of relevant outcomes, while recall is the proportion of relevant results accurately classified by the network. To simplify, another statistic, the F-1 score, was utilized. In the final step, a dataset not included in the training dataset was used for testing. This work utilized Keras ("Keras. models. Sequential"), Python 3.5, and the Numpy, SKLearn, and Scipy library packages to implement the FCN model. Convolutional layers, also known as Keras layers and Convolutional2D layers, as well as fully connected layers, were used in building the neural network (Keras. layers. Dense).

3. Results

The Fully Convolutional Network model was fed the input Kitti lane dataset. The network's performance is demonstrated in this section, and the graphical results are also depicted. Subsequently, the effectiveness of the models using the Kitti lane dataset was evaluated. Note that all findings were reviewed using the dataset of validation, as the Kitti lane dataset website only permits the use of test data for reporting purposes. In addition, this section lists and describes the results of the complete version on the benchmark server compared to other common approaches. In addition, the FCN model for lane detection that employs an efficient channel attention mechanism is known as the FCN-with efficient channel attention.

In contrast, the FCN model that employs the channel attention mechanism for the lane detection task is called the FCN-with channel attention. Table 3 provides an overview of four input model metrics that use the KITTI validation set. These metrics are the F1-measure, accuracy, precision, and recall. According to the quantitative results, the suggested FCN-with efficient channel attention mechanism slightly outperformed the FCN-with channel attention mechanism across the four criteria. Table 3 also reveals that the training accuracy of the FCN model with efficient channel attention was marginally higher than that of the FCN model with channel attention, at 96.13 percent. The FCN with channel attention received a result of 96 percent of accuracy.

Table 3
Setting of parameters

Method	Accuracy	F1-measure	Precision	Recall
FCN-with efficient channel attention	96.13	96.08	96.14	96.06
FCN-with channel attention	96.00	95.39	95.48	95.31

Next, Figure 3 displays the Kitti lane dataset's visualization results. The green lines represent the detected lane markings. These findings suggest that the suggested FCN-with efficient channel attention and FCN-with channel attention are effective at lane detection. The Fully Convolutional Network (FCN) predicts comparable outputs from any size and randomness inputs. Using dense-feed forward computing and backpropagation, it executes a straightforward and coherent learning procedure from an entire input dataset concurrently with the ground truths. In addition, promising accuracy was achieved by transferring pre-trained weights, combining distinct layer representations, and learning the end-to-end of entire images. In addition, due to the integration of FCN with an attention mechanism, the FCN was able to acquire a greater understanding of image features. Therefore, combining FCN with an attention mechanism improved accuracy by increasing the layers' capability to represent features, enhancing feature localization in the feature map, emphasizing critical spatial information, and possibly increasing the weighted information of lane line targets while suppressing irrelevant information.

Furthermore, the attention method employed to fuse information improves the model's robustness. As an additional perk, it boosted the network's performance even further without adding any new annotations. In addition, it does not lengthen the inference time. This study adds the attention mechanism to significantly improve the network's ability to extract features. Therefore, the application of attention enhances the localization of features on the feature map. Combining the FCN model with the attention mechanism improved the lane detection model's performance, allowing for reliable lane detection, illumination circumstances, and road morphologies.

(a)

(b)

Fig. 3. Results of the visualization of the Kitti lane dataset validated in the (a) FCN-with efficient channel attention model and (b) FCN-with channel attention model

Next, the proposed FCN-with efficient channel attention mechanism and FCN-with channel attention mechanism detection techniques were compared with other popular methods demonstrated in the KITTI lane dataset website. Table 4 presents the rankings achieved by the top ten real-name algorithms competing in the urban road category of the KITTI Vision Benchmark Suite Server compared to the suggested technique. These results are displayed on the leaderboard of the benchmark. The listed methods, namely RBANet, DGIST MT-CNN, RBNet, MultiNet, Hadamard-FCN, LoDNN, DEEP-DIG, FTP, FCN-LC, and MAP are FCN-based deep learning methods approach that utilizing Kitti lane dataset to evaluate their proposed methods for lane detection in autonomous driving [50-58].

The RBANet integrated the Segnet method with the boundary and reverse attention mechanisms to provide additional context. In the meantime, DGIST MT-CNN applied the multi-task deep learning technique and employed a complete convolution CNN (one-stage) model-based 2D object identification model. Meanwhile, RBNet carried out the road detection using the Res50 pre-trained model. Next, an improved deconvolution strategy for MultiNet has been proposed. This strategy uses VGG as the encoder and FCN as the decoder. Hadamard-FCN is a fully convolutional network variation that adds additional element-wise layers. It is based on the Inception v2 feature encoder and uses it as its foundation. In the meantime, LoDNN is employing FCN for road segmentation. However, rather than the RGB data utilized by other approaches, LoDNN uses the top-view LIDAR images.

After that, DEEP-DIG methods utilized the ResNet-101 FCN with three skip connections, four upsampling steps, and bilinear kernels. Next, the model was first trained on ImageNet. Then it was fine-tuned using random data augmentation on an image of the KITTI road training set in Bird's-Eye View that included geometric modifications and pixel value changes. Next, we have the FTP method, which uses a Finetuning Deeplab FCN pre-trained on Imagenet. FCN-LC, on the other hand, is a convolutional ANN that has been trained on patches, and inference is carried out with the help of a fully convolutional network. Lastly, MAP is a method that projects the OpenStreetMap data onto the image plane and then refines the projected labels onto the image. This generates training labels for the pixels in the image. These labels were created automatically to train the FCN.

The suggested technique obtained a competitive F1 measure of 96%, which differed marginally from RBANet. As a result, the proposed model's results are in the top three, and these methods can compete with other significant algorithms in general. Furthermore, according to Table 4, compared to the different approaches under FCN, the suggested method can obtain the best results for all criteria (F1-measure, Precision, and Recall).

Table 4
The leader board of the Kitti Vision Benchmark Suite Server's list of the leading ten real-name algorithms for the urban road category (in %)

Method	F1-measure	Precision	Recall
FCN-with efficient channel attention	96.08	96.14	96.06
FCN-with channel attention	95.39	95.48	95.31
RBANet [51]	96.30	95.14	97.50
DGIST MT-CNN [52]	95.60	95.84	95.52
RBNet [53]	94.97	94.94	95.01
MultiNet [54]	94.88	94.84	94.91
Hadamard-FCN [55]	94.85	94.81	94.89
LoDNN [50]	94.07	92.81	95.37
DEEP-DIG [56]	93.98	94.26	93.69
FTP [57]	91.61	91.04	92.20
FCN-LC [58]	90.79	90.87	90.72
MAP [57]	87.80	86.01	89.66

In addition, the suggested network can achieve a high accuracy level since the attention mechanism enhances the performance of deep models. The outcome reveals that the attention mechanism can provide higher-level characteristics than conventional FCN. In addition, the deployment of the FCN with attention integration demonstrates great accuracy and reduced attention loss because:

i. Incorporating attention into the neural network increases the localization of image features in the feature map.
ii. Increase feature representation on thin and sparse lane marking annotations.
iii. Enhance the feature extraction and feature representation capabilities of the network without increasing inference time.

4. Conclusions

GmbH, which was authorized for its creation by Jannik Fritsch and Tobias Kuehl. The dataset had 579 images with a resolution of 375 by 1242 pixels, including 289 training images alongside their ground truth and 290 test images. After that, the proposed model that presents two different configurations of the attention mechanism was developed for lane detection. These configurations are known as FCN-with efficient channel attention and FCN-with channel attention, respectively. Both models achieved an accuracy of greater than 95%, which provides cause for optimism about their performance. As a result, the proposed method has the potential to learn additional discriminative qualities of the road lanes. Furthermore, the results demonstrated that the introduced model contributed a high-level feature to the network, making it possible for a single network model to achieve performance that is still competitive within the detecting lanes. A comparison of the two approaches suggested found that the FCN-efficient channel attention performed significantly better. In the future, various other deep learning methods, such as the Recurrent Neural Network, the Faster R-CNN, and the Convolutional Neural Network (CNN), could be used to analyze the various structures of input features to improve the performance of lane detection. This would be done to improve the performance of lane detection that incorporates a variety of attention mechanisms. In addition, this would be performed to enhance the method's accuracy.

Acknowledgement
This research was funded by a grant from Ministry of Higher Education (MOHE) through Fundamental Research Grant Scheme (FRGS), Registration Proposal No: FRGG/1/2022/ICT11/UTM/02/2 (Attention-Based Fully Convolutional Networks for Lane Detection of Different Driving Scenes - No Vote: R.K130000.7843.5F563).

References
[1] Khairi, Danial Mohd, Mohd Azman Abas, Mohd Farid Muhamad Said, and Wan Saiful-Islam Wan Salim. "Fuel consumption mathematical models for road vehicle–A review." *Progress in Energy and Environment* (2021): 59-71.
[2] Zhang, Lu, Fengling Jiang, Bin Kong, Jing Yang, and Can Wang. "Real-Time Lane Detection by Using Biologically Inspired Attention Mechanism to Learn Contextual Information." *Cognitive Computation* 13 (2021): 1333-1344. https://doi.org/10.1007/s12559-021-09935-5
[3] Zhang, Ronghui, Yueying Wu, Wanting Gou, and Junzhou Chen. "RS-lane: A robust lane detection method based on Resnest and self-attention distillation for challenging traffic situations." *Journal of advanced transportation* 2021 (2021): 1-12. https://doi.org/10.1155/2021/7544355
[4] Munir, Farzeen, Shoaib Azam, Moongu Jeon, Byung-Geun Lee, and Witold Pedrycz. "LDNet: End-to-end lane marking detection approach using a dynamic vision sensor." *IEEE Transactions on Intelligent Transportation Systems* 23, no. 7 (2021): 9318-9334. https://doi.org/10.1109/TITS.2021.3102479
[5] Li, Mingfa, Yuanyuan Li, and Min Jiang. "Lane detection based on connection of various feature extraction methods." *Advances in Multimedia* 2018 (2018). https://doi.org/10.1155/2018/8320207
[6] Huang, Yuhao, Shitao Chen, Yu Chen, Zhiqiang Jian, and Nanning Zheng. "Spatial-temproal based lane detection using deep learning." In *Artificial Intelligence Applications and Innovations: 14th IFIP WG 12.5 International Conference, AIAI 2018, Rhodes, Greece, May 25–27, 2018, Proceedings 14*, pp. 143-154. Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-92007-8_13
[7] Devane, Vighnesh, Ganesh Sahane, Hritish Khairmode, and Gaurav Datkhile. "Lane detection techniques using image processing." In *ITM Web of Conferences*, vol. 40, p. 03011. EDP Sciences, 2021. https://doi.org/10.1051/itmconf/20214003011
[8] Kodeeswari, M., and Philemon Daniel. "Lane line detection in real time based on morphological operations for driver assistance system." In *2017 4th International Conference on Signal Processing, Computing and Control (ISPCC)*, pp. 316-320. IEEE, 2017. https://doi.org/10.1109/ISPCC.2017.8269696
[9] Kavitha, D., and S. Ravikumar. "Designing an IoT based autonomous vehicle meant for detecting speed bumps and lanes on roads." *Journal of Ambient Intelligence and Humanized Computing* 12 (2021): 7417-7426. https://doi.org/10.1007/s12652-020-02419-8
[10] Gong, Jinliang, Tao Chen, and Yanfei Zhang. "Complex lane detection based on dynamic constraint of the double threshold." *Multimedia Tools and Applications* 80, no. 18 (2021): 27095-27113. https://doi.org/10.1007/s11042-021-10978-x
[11] Akbari, Behzad, Jeyan Thiyagalingam, Richard Lee, and Kirubarajan Thia. "A multilane tracking algorithm using IPDA with intensity feature." *Sensors* 21, no. 2 (2021): 461. https://doi.org/10.3390/s21020461
[12] Saussard, Romain, Boubker Bouzid, Marius Vasiliu, and Roger Reynaud. "The embeddability of lane detection algorithms on heterogeneous architectures." In *2015 IEEE International Conference on Image Processing (ICIP)*, pp. 4694-4697. IEEE, 2015. https://doi.org/10.1109/ICIP.2015.7351697

[13] Yousri, Retaj, Mustafa A. Elattar, and M. Saeed Darweesh. "A deep learning-based benchmarking framework for lane segmentation in the complex and dynamic road scenes." *IEEE Access* 9 (2021): 117565-117580. https://doi.org/10.1109/ACCESS.2021.3106377

[14] Von Gioi, Rafael Grompone, Jérémie Jakubowicz, Jean-Michel Morel, and Gregory Randall. "LSD: A line segment detector." *Image Processing On Line* 2 (2012): 35-55. https://doi.org/10.5201/ipol.2012.gjmr-lsd

[15] Hota, Rudra N., Shahanaz Syed, Subhadip Bandyopadhyay, and P. Radha Krishna. "A Simple and Efficient Lane Detection using Clustering and Weighted Regression." In *COMAD*. 2009.

[16] Hu, Jianjun, Songsong Xiong, Yuqi Sun, Junlin Zha, and Chunyun Fu. "Research on lane detection based on global search of dynamic region of interest (DROI)." *Applied Sciences* 10, no. 7 (2020): 2543. https://doi.org/10.3390/app10072543

[17] Tabelini, Lucas, Rodrigo Berriel, Thiago M. Paixao, Claudine Badue, Alberto F. De Souza, and Thiago Oliveira-Santos. "Polylanenet: Lane estimation via deep polynomial regression." In *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 6150-6156. IEEE, 2021. https://doi.org/10.1109/ICPR48806.2021.9412265

[18] Li, Wenhui, Feng Qu, Ying Wang, Lei Wang, and Yuhao Chen. "A robust lane detection method based on hyperbolic model." *Soft Computing* 23 (2019): 9161-9174. https://doi.org/10.1007/s00500-018-3607-x

[19] Kasmi, Abderrahim, Johann Laconte, Romuald Aufrère, Ruddy Theodose, Dieumet Denis, and Rol Chapuis. "An information driven approach for ego-lane detection using Lidar and OpenStreetMap." In *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 522-528. IEEE, 2020. https://doi.org/10.1109/ICARCV50220.2020.9305388

[20] Dorj, Byambaa, Sabir Hossain, and Deok-Jin Lee. "Highly curved lane detection algorithms based on Kalman filter." *Applied Sciences* 10, no. 7 (2020): 2372. https://doi.org/10.3390/app10072372

[21] Ghanem, Safwan, Priyadarshi Kanungo, Ganapati Panda, Suresh Chandra Satapathy, and Rohit Sharma. "Lane detection under artificial colored light in tunnels and on highways: an IoT-based framework for smart city infrastructure." *Complex & Intelligent Systems* (2021): 1-12. https://doi.org/10.1007/s40747-021-00381-2

[22] I, Robot. "Artificial Intelligence." In *Philosophy through Film*, pp. 102-129. Routledge, 2019. https://doi.org/10.4324/9780203772294-10

[23] Alzubaidi, Laith, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, José Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions." *Journal of big Data* 8 (2021): 1-74. https://doi.org/10.1186/s40537-021-00444-8

[24] Wu, Zhenyu, Kai Qiu, Tingning Yuan, and Hongmei Chen. "A method to keep autonomous vehicles steadily drive based on lane detection." *International Journal of Advanced Robotic Systems* 18, no. 2 (2021): 17298814211002974. https://doi.org/10.1177/17298814211002974

[25] Sun, Yuxiang, Lujia Wang, Yongquan Chen, and Ming Liu. "Accurate lane detection with atrous convolution and spatial pyramid pooling for autonomous driving." In *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 642-647. IEEE, 2019. https://doi.org/10.1109/ROBIO49542.2019.8961705

[26] Philion, Jonah. "Fastdraw: Addressing the long tail of lane detection by adapting a sequential prediction network." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11582-11591. 2019. https://doi.org/10.1109/CVPR.2019.01185

[27] Muthalagu, Raja, Anudeepsekhar Bolimera, and V. Kalaichelvi. "Vehicle lane markings segmentation and keypoint determination using deep convolutional neural networks." *Multimedia Tools and Applications* 80, no. 7 (2021): 11201-11215. https://doi.org/10.1007/s11042-020-10248-2

[28] Alajlan, Abrar Mohammed, and Marwah Mohammad Almasri. "Automatic lane marking prediction using convolutional neural network and S-Shaped Binary Butterfly Optimization." *The Journal of Supercomputing* 78, no. 3 (2022): 3715-3745. https://doi.org/10.1007/s11227-021-03988-x

[29] Kanagaraj, Nitin, David Hicks, Ayush Goyal, Sanju Tiwari, and Ghanapriya Singh. "Deep learning using computer vision in self driving cars for lane and traffic sign detection." *International Journal of System Assurance Engineering and Management* 12, no. 6 (2021): 1011-1025. https://doi.org/10.1007/s13198-021-01127-6

[30] Zhan, Huiqin, and Lin Chen. "Lane detection image processing algorithm based on FPGA for intelligent vehicle." In *2019 Chinese Automation Congress (CAC)*, pp. 1190-1196. IEEE, 2019. https://doi.org/10.1109/CAC48633.2019.8996283

[31] Ye, Yang Yang, Xiao Li Hao, and Hou Jin Chen. "Lane detection method based on lane structural analysis and CNNs." *IET Intelligent Transport Systems* 12, no. 6 (2018): 513-520. https://doi.org/10.1049/iet-its.2017.0143

[32] Zhang, Ge, Chaokun Yan, and Jianlin Wang. "Quality-guided lane detection by deeply modeling sophisticated traffic context." *Signal Processing: Image Communication* 84 (2020): 115811. https://doi.org/10.1016/j.image.2020.115811

[33] Feng, Zhaofei, Shuo Zhang, Martin Kunert, and Werner Wiesbeck. "Applying neural networks with a high-resolution automotive radar for lane detection." In *AmE 2019-Automotive meets Electronics; 10th GMM-Symposium*, pp. 1-6. VDE, 2019.
[34] Liu, Jingyi. "Learning full-reference quality-guided discriminative gradient cues for lane detection based on neural networks." *Journal of Visual Communication and Image Representation* 65 (2019): 102675. https://doi.org/10.1016/j.jvcir.2019.102675
[35] Zou, Qin, Hanwen Jiang, Qiyu Dai, Yuanhao Yue, Long Chen, and Qian Wang. "Robust lane detection from continuous driving scenes using deep neural networks." *IEEE transactions on vehicular technology* 69, no. 1 (2019): 41-54. https://doi.org/10.1109/TVT.2019.2949603
[36] Pihlak, René, and Andri Riid. "Simultaneous road edge and road surface markings detection using convolutional neural networks." In *International Baltic Conference on Databases and Information Systems*, pp. 109-121. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-57672-1_9
[37] Chng, Zhe Ming, Joseph Mun Hung Lew, and Jimmy Addison Lee. "RONELD: Robust neural network output enhancement for active lane detection." In *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 6842-6849. IEEE, 2021. https://doi.org/10.1109/ICPR48806.2021.9412572
[38] Pizzati, Fabio, Marco Allodi, Alejandro Barrera, and Fernando García. "Lane detection and classification using cascaded CNNs." In *Computer Aided Systems Theory–EUROCAST 2019: 17th International Conference, Las Palmas de Gran Canaria, Spain, February 17–22, 2019, Revised Selected Papers, Part II 17*, pp. 95-103. Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-45096-0_12
[39] Zhang, Lu, Fengling Jiang, Bin Kong, Jing Yang, and Can Wang. "Real-Time Lane Detection by Using Biologically Inspired Attention Mechanism to Learn Contextual Information." *Cognitive Computation* 13 (2021): 1333-1344. https://doi.org/10.1007/s12559-021-09935-5
[40] Li, Jingyu, Fengling Jiang, Jing Yang, Bin Kong, Mandar Gogate, Kia Dashtipour, and Amir Hussain. "Lane-deeplab: Lane semantic segmentation in automatic driving scenarios for high-definition maps." *Neurocomputing* 465 (2021): 15-25. https://doi.org/10.1016/j.neucom.2021.08.105
[41] Hernández, Adrián, and José M. Amigó. "Attention mechanisms and their applications to complex systems." *Entropy* 23, no. 3 (2021): 283. https://doi.org/10.3390/e23030283
[42] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132-7141. 2018. https://doi.org/10.1109/CVPR.2018.00745
[43] Chang, Donghoon, Vinjohn Chirakkal, Shubham Goswami, Munawar Hasan, Taekwon Jung, Jinkeon Kang, Seok-Cheol Kee, Dongkyu Lee, and Ajit Pratap Singh. "Multi-lane detection using instance segmentation and attentive voting." In *2019 19th International conference on control, automation and systems (ICCAS)*, pp. 1538-1542. IEEE, 2019. https://doi.org/10.23919/ICCAS47443.2019.8971488
[44] Aly, Mohamed. "Real time detection of lane markers in urban streets." In *2008 IEEE intelligent vehicles symposium*, pp. 7-12. IEEE, 2008. https://doi.org/10.1109/TITS.2017.2723575
[45] Cordts, Marius, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. "The cityscapes dataset for semantic urban scene understanding." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3213-3223. 2016. https://doi.org/10.1109/CVPR.2016.350
[46] Fritsch, Jannik, Tobias Kuehnl, and Andreas Geiger. "A new performance measure and evaluation benchmark for road detection algorithms." In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pp. 1693-1700. IEEE, 2013. https://doi.org/10.1109/ITSC.2013.6728473
[47] Zheng, Zhuo, Yanfei Zhong, Ailong Ma, and Liangpei Zhang. "FPGA: Fast patch-free global learning framework for fully end-to-end hyperspectral image classification." *IEEE Transactions on Geoscience and Remote Sensing* 58, no. 8 (2020): 5612-5626. https://doi.org/10.1109/TGRS.2020.2967821
[48] Guo, Meng-Hao, Tian-Xing Xu, Jiang-Jiang Liu, Zheng-Ning Liu, Peng-Tao Jiang, Tai-Jiang Mu, Song-Hai Zhang, Ralph R. Martin, Ming-Ming Cheng, and Shi-Min Hu. "Attention mechanisms in computer vision: A survey." *Computational visual media* 8, no. 3 (2022): 331-368. https://doi.org/10.1007/s41095-022-0271-y
[49] Wang, Jianqing, Weitao Mo, Yan Wu, Xiaomei Xu, Yi Li, Jianming Ye, and Xiaobo Lai. "Combined Channel Attention and Spatial Attention Module Network for Chinese Herbal Slices Automated Recognition." *Frontiers in Neuroscience* 16 (2022): 920820. https://doi.org/10.3389/fnins.2022.920820
[50] Caltagirone, Luca, Samuel Scheidegger, Lennart Svensson, and Mattias Wahde. "Fast LIDAR-based road detection using fully convolutional neural networks." In *2017 ieee intelligent vehicles symposium (iv)*, pp. 1019-1024. IEEE, 2017. https://doi.org/10.1109/IVS.2017.7995848
[51] Sun, Jee-Young, Seung-Wook Kim, Sang-Won Lee, Ye-Won Kim, and Sung-Jea Ko. "Reverse and boundary attention network for road segmentation." In *Proceedings of the IEEE/CVF international conference on computer vision workshops*, pp. 0-0. 2019. https://doi.org/10.1109/ICCVW.2019.00116

[52] Kumar, G. Ajay, Jin Hee Lee, Jongrak Hwang, Jaehyeong Park, Sung Hoon Youn, and Soon Kwon. "LiDAR and camera fusion approach for object distance estimation in self-driving vehicles." *Symmetry* 12, no. 2 (2020): 324. https://doi.org/10.3390/sym12020324

[53] Chen, Zhe, and Zijing Chen. "Rbnet: A deep neural network for unified road and road boundary detection." In *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part I 24*, pp. 677-687. Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-70087-8_70

[54] Teichmann, Marvin, Michael Weber, Marius Zoellner, Roberto Cipolla, and Raquel Urtasun. "Multinet: Real-time joint semantic reasoning for autonomous driving." In *2018 IEEE intelligent vehicles symposium (IV)*, pp. 1013-1020. IEEE, 2018. https://doi.org/10.1109/IVS.2018.8500504

[55] Oeljeklaus, Malte. "An integrated approach for traffic scene understanding from monocular cameras." (2020). https://doi.org/10.51202/9783186815125

[56] Munoz-Bulnes, Jesús, Carlos Fernandez, Ignacio Parra, David Fernández-Llorca, and Miguel A. Sotelo. "Deep fully convolutional networks with random data augmentation for enhanced generalization in road detection." In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 366-371. IEEE, 2017. https://doi.org/10.1109/ITSC.2017.8317901

[57] Laddha, Ankit, Mehmet Kemal Kocamaz, Luis E. Navarro-Serment, and Martial Hebert. "Map-supervised road detection." In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pp. 118-123. IEEE, 2016. https://doi.org/10.1109/IVS.2016.7535374

[58] Mendes, Caio César Teodoro, Vincent Frémont, and Denis Fernando Wolf. "Exploiting fully convolutional neural networks for fast road detection." In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3174-3179. IEEE, 2016. https://doi.org/10.1109/ICRA.2016.7487486