International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 66-73 www.ijaem.net ISSN: 2395-5252

A Novel Lane Detection System Using Enhanced Multi-Lane Parameter Optimization

A. Milcah Nimalin
PG Scholars, Department of Computer Science and Engineering, C.S.I Institute of Technology, Thovalai, Tamilnadu, India

Submitted: 01-08-2022 Revised: 02-08-2022 Accepted: 08-08-2022

ABSTRACT
Lane detection is a challenging problem. It has attracted the attention of the computer vision community for several decades. Essentially, lane detection is a multi-feature detection problem that has become a real challenge for computer vision and machine learning techniques. Although many machine learning methods are used for lane detection, they are mainly used for classification rather than feature design. But modern machine learning methods can be used to identify the features that are rich in recognition and have achieved success in feature detection tests. With increase in the number of road accidents, it has led to concern over the nature of accidents. Most of the time, it is due to human error. So the lane detection systems are being developed for assisting the driver. The main purpose of it is to detect the lanes and warn the driver of lane departure. The proposed system is a novel lane detection system, called Scene Understanding Physics-Enhanced Real-time (SUPER) algorithm. The proposed method consists of two main modules: Hierarchical Semantic Segmentation network as the scene feature extractor and a physics enhanced multi-lane parameter optimization module for lane inference. It includes an optimization framework to estimate lane parameters. The hierarchical structure extracts the objects in each level and this structure tries to label an image of a street scene into coarse geometric classes that are useful for tasks such as navigation, object recognition, and general scene understanding. Assuming that lane markings are largely parallel polynomials, it separate the lane parameters into shared parts (heading angle and curvature) and unique parts (offsets and lane marking attributes). To cope with non-flat ground, a polynomial road model is adopted. Then the lane parameter estimation problem is formulated as an optimization problem based on the pixel-wise scene labels, solving which can efficiently estimate all parameters.

Keywords: Lane detection (LD), Scene understanding, Semantic segmentation, Lane markings.

I. INTRODUCTION

Lane detection is a multi-feature detection problem that has become a real challenge for computer vision and machine learning techniques. The major principal approaches is to detect road boundaries and lanes using vision system on the vehicle. However, due to the diverse appearances arising from adverse lighting/weather conditions and presence of other objects, lane detection is still a challenging problem. Most methods remain riddled with assumptions and limitations, still not good enough for safe and reliable driving in the real world. Lane detection can be achieved by using monocular cameras, stereo cameras, lidars, etc. Cameras are most popular due to their rich content features and affordable cost. Many feature-based methods use the generic framework. They decompose road/lane detection methods into several modules: image pre-processing, feature extraction, model fitting, and image to world correspondence and time integration.

II. LANE DETECTION (LD)

Lane detection is the task of detecting lanes on a road from a camera. Lane detection is an important foundation in the development of intelligent vehicles. The proposed algorithm is crucial in promoting the technological level of intelligent vehicle driving assistance and conducive to the further improvement of the driving safety of intelligent vehicles.

1. WHAT IS FRAME CONVERSION IN IMAGE PROCESSING?

Frames can be obtained from a video and

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 66-73 www.ijaem.net ISSN: 2395-5252

converted into images. A video is a sequence of images (called frames) captured and eventually displayed at a given frequency. However, by stopping at a specific frame of the sequence, a single video frame, i.e. an image, is obtained. It is the presentation of visual elements in an image, especially the placement of the subject in relation to other objects. It can make an image more aesthetically pleasing and keep the viewer's focus on the framed object.

2. WHAT IS CONVOLUTION IN IMAGE PROCESSING?

Convolution is the process of transforming an image by applying a kernel over each pixel and its local neighbors across the entire image. The kernel is a matrix of values whose size and values determine the transformation effect of the convolution process. The convolution layers are used to help the computer determine features that could be missed in simply flattening an image into its pixel values. These filters can be to highlight simple features, such as vertical or horizontal lines to make it more obvious to the computer what it is looking at. It has applications that include probability, statistics, acoustics, spectroscopy, signal processing and image processing, geophysics, engineering, physics, computer vision and differential equations. It can be used to modify the image (e.g. blurring), find relevant structures (e.g. edge detection) or infer arbitrary features.

3. WHAT IS HIERARICHICAL SEMANTIC SEGMENTATION?

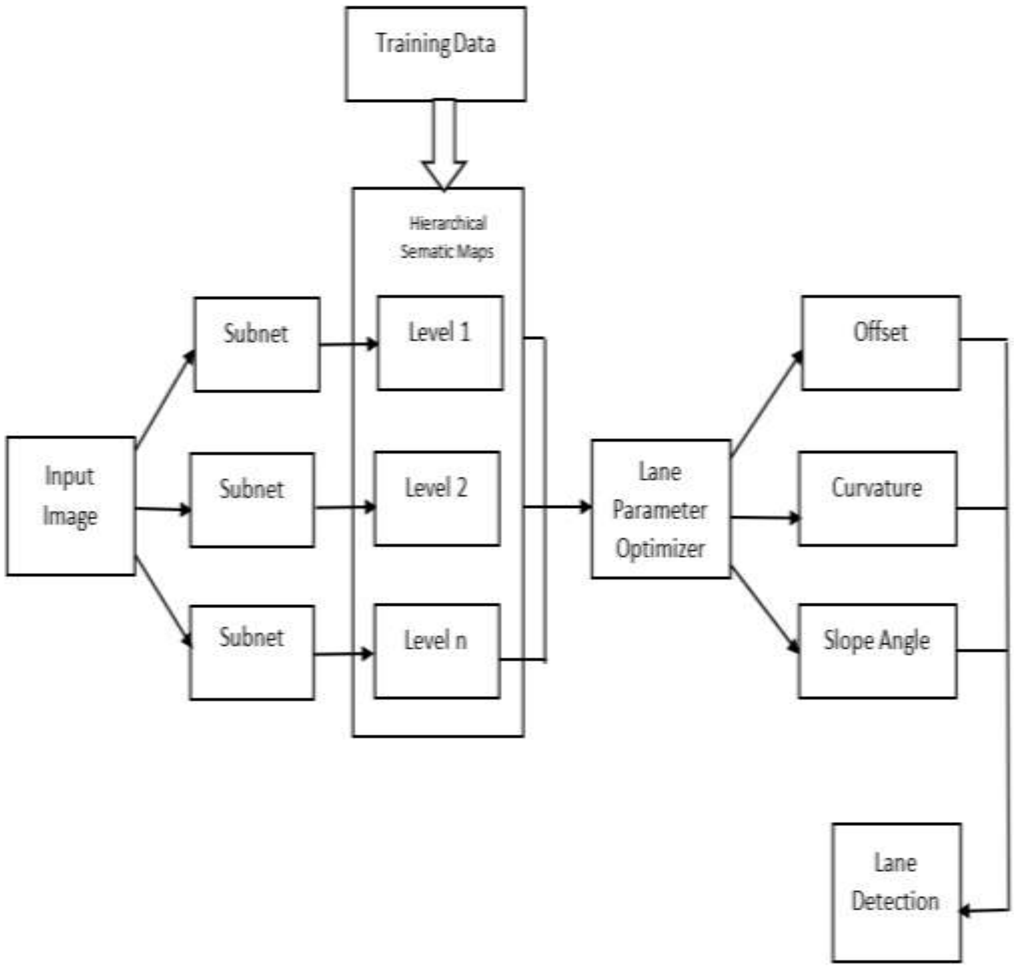
Semantic Segmentation is the process of assigning a label to every pixel in the image. This is in stark contrast to classification, where a single label is assigned to the entire picture. Semantic segmentation treats multiple objects of the same class as a single entity. It can reduce the complexity of the image, and thus analyzing the image becomes simpler. Segmentation is typically used to locate objects and boundaries in images. It is an important stage of the image recognition system, because it extracts the objects of our interest, for further processing such as description or recognition. Segmentation techniques are used to isolate the desired object from the image in order to perform analysis of the object.

III. SYSTEM ARCHITECTURE

The architectural design describes the overall representation of the proposed system Convolution performs blurring, sharpening, edge detection, noise reduction etc., Deconvolution is a computationally intensive image processing technique that is being increasingly utilized for improving the contrast and resolution of digital images captured in the microscope. Bilinear upsampling is the process of creating a larger resolution image where every sample is created from bilinear filtering of a smaller resolution image. The core idea is that CNNs are used for scene understanding as well as road/lane extraction, while a physical road/lane models are adopted for the lane inference. The explain ability is improved, and the subsequent lane inference module that uses a collaborative optimization can be more accurate and reliable. An improved hierarchical semantic segmentation structure to capture lane related information focusing on the region of interest, i.e., where lane markings are likely to exist, and also enable training on multiple heterogeneous datasets. An optimization-based lane inference method to directly estimate all parameters (i.e., lateral offsets, heading angle, and curvature) of multiple lanes in real-time.

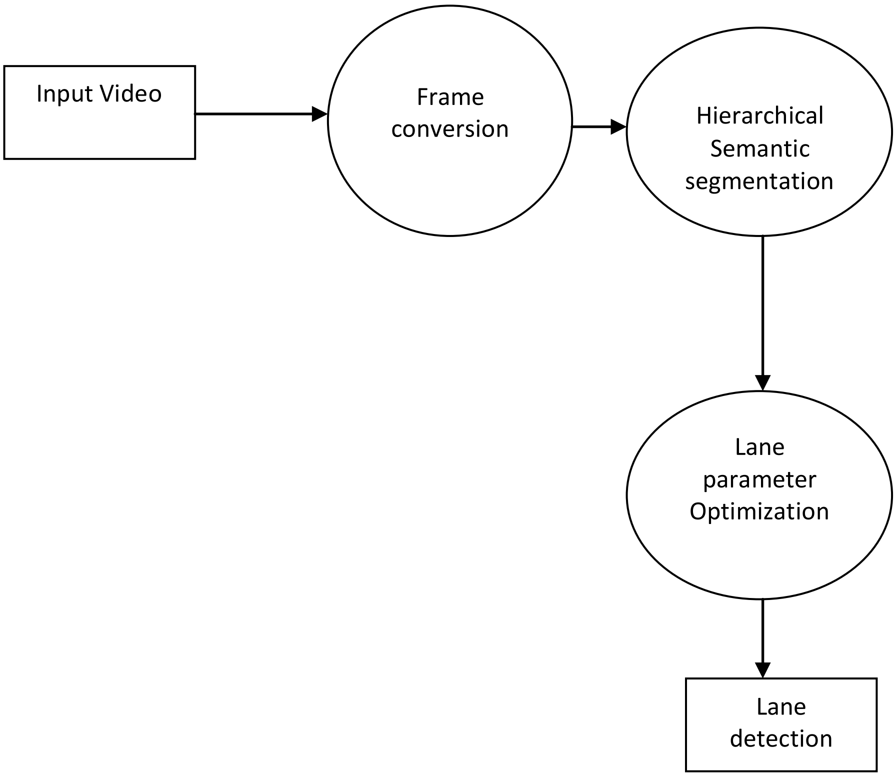
A hierarchical semantic segmentation module and a physics enhanced multi-lane inference module. The core idea is that CNNs are used for scene understanding as well as road/lane extraction, while a physical road/lane models are adopted for the lane inference. The lane inference module that uses a collaborative optimization can be more accurate and reliable. An improved hierarchical semantic segmentation structure to capture lane related information focusing on the region of interest, i.e., where lane markings are likely to exist, and also enable training on multiple heterogeneous datasets. An optimization-based lane inference method to directly estimate all parameters (i.e., lateral offsets, heading angle, and curvature) of multiple lanes in real-time. Multi-level classification is obtained. Four-level classifiers are used.

- First level - an image is divided into sky, vertical and support areas;
- Second level - vertical/support areas are further divided into
- Third level - drivable areas are further classified;
- Last level - detailed lane marking types are extracted from the lane marking areas.



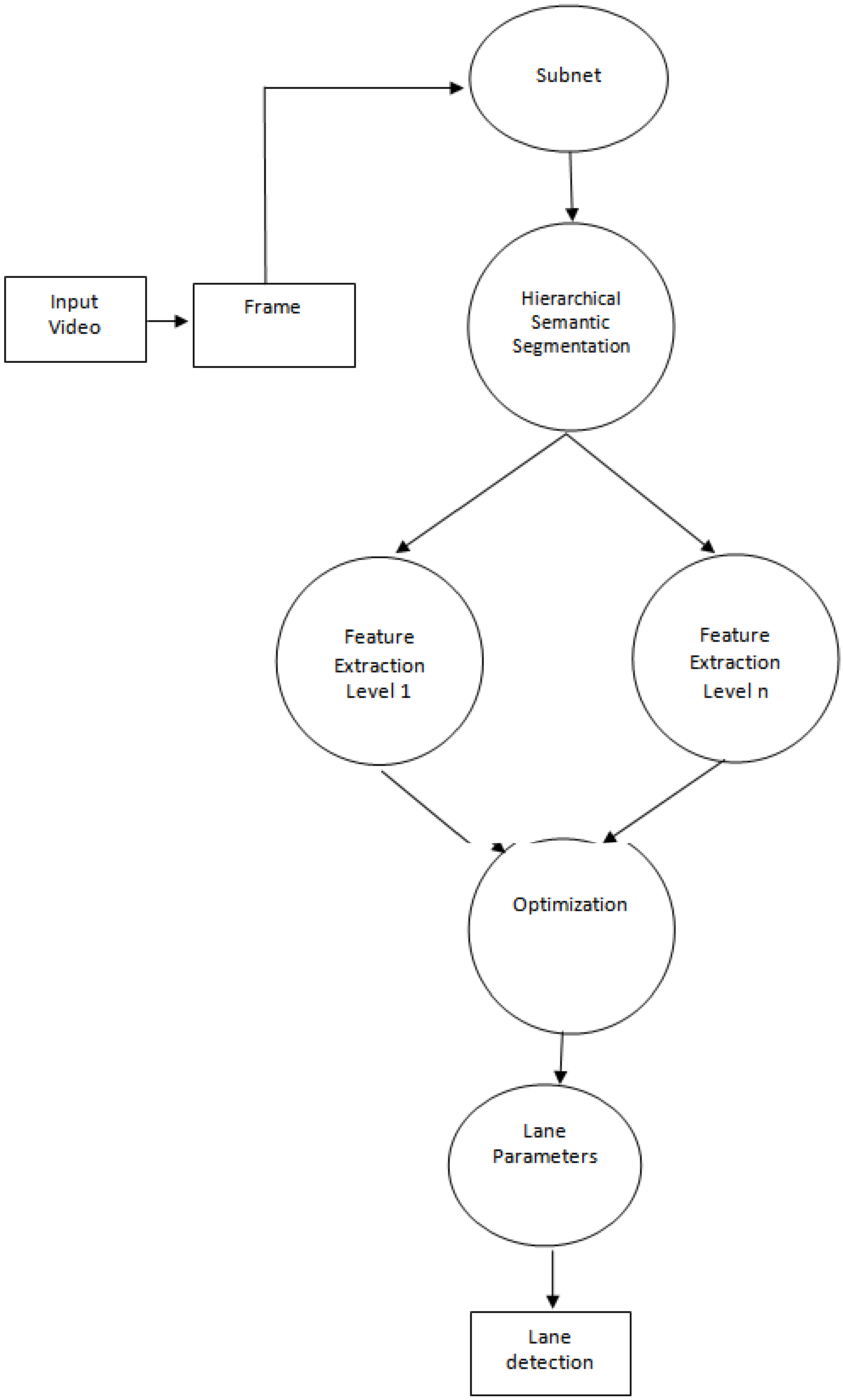
Level 0 DFD

The Level 0 DFD describes the overall data flow of the proposed system. In this the input video is the input data and the process such as frame conversion, Semantic segmentation and optimization process is obtained to perform the lane detection process. The proposed hierarchical structure extracts the objects in each level and this structure tries to label an image of a street scene into coarse geometric classes that are useful for tasks such as navigation, object recognition, and general scene understanding. The four-level classifiers are designed to perform the segmentation process. In the first level, an image is divided into sky, vertical and support areas; in the second level, vertical/support areas are further divided into subclasses for scene description; in the third level, drivable areas are further classified; in the last level, detailed lane marking types are extracted from the lane marking areas. Once scene labels are predicted by the hierarchical semantic segmentation, this module can use lane related labels to estimate lane parameters in the vehicle coordinate system. It assume the lane lines are largely parallel polynomials, and the lane parameters can be divided into two parts, shared/global parameters (i.e., heading angle and curvature) and unique/local parameters (i.e., offset of each lane). The multi-lane parameters of road slope is proposed for better accuracy. The Nelder-Mead Simplex (NMS) algorithm is used for lane detection. NMS maintains a simplex, and extrapolates the behavior of the cost function to find a new test point to replace the worst point in a simplex through deterministic rules, i.e., reflection, expansion, contraction, and shrink. The above simplex evolution procedure will repeat until convergence. Usually, a proper initial guess helps to avoid trapped at the local minima and speed up the searching process. The curvature information of the road ahead the vehicle from map or navigation software can also be used. If no prior information is available, a proper range is selected for each parameter and the initial values are set to zero.



Level 1 DFD

The Level 1 DFD describes the detailed data flow of the proposed system. . In this the input video is the input data and the process such as frame conversion, Semantic segmentation and optimization process is obtained to perform the lane detection process. The semantic segmentation process the frames to extract the image feature at different levels. From the extracted features the lane parameters are identified and optimized to perform the effective lane detection process. CNNs are used for scene understanding as well as road/lane extraction, while a physical road/lane models are adopted for the lane inference. The explain ability is improved, and the subsequent lane inference module that uses a collaborative optimization can be more accurate and reliable. An improved hierarchical semantic segmentation structure to capture lane related information focusing on the region of interest, i.e., where lane markings are likely to exist, and also enable training on multiple heterogeneous datasets. An optimization-based lane inference method to directly estimate all parameters (i.e., lateral offsets, heading angle, and curvature) of multiple lanes in real-time. Especially, an ingeniously designed cost function considering road/lane geometric models is proposed to formulate the parameter estimation problem. Once scene labels are predicted by the hierarchical semantic segmentation, this module can use lane related labels to estimate lane parameters in the vehicle coordinate system. It assume the lane lines are largely parallel polynomials, and the lane parameters can be divided into two parts, shared/global parameters (i.e., heading angle and curvature) and unique/local parameters (i.e., offset of each lane). Other lane marking attributes such as type and color are detected separately. The multi-lane parameters of road slope is proposed for better accuracy. The lane parameter estimation is considered as an optimization problem. The cost function is designed to reflect the physical properties of lane lines.


Subnet
Input Video
Frame
Hierarchical Semantic Segmentation
Feature Extraction Level 1
Feature Extraction Level n
Optimization
Lane Parameters
Lane detection

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 66-73 www.ijaem.net ISSN: 2395-5252

IV. SYSTEM IMPLEMENTATION

A novel lane detection system named as Scene Understanding Physics-Enhanced Real-time (SUPER). Its main difference from existing methods is that it solve the problem starting from scene understanding, not focusing on lane markings only, and then estimate lane parameters through optimization with a physics-based cost function. A hierarchical semantic segmentation module and a physics enhanced multi-lane inference module.

The core idea is that CNNs are used for scene understanding as well as road/lane extraction, while a physical road/lane models are adopted for the lane inference. The explain ability is improved, and the subsequent lane inference module that uses a collaborative optimization can be more accurate and reliable. An improved hierarchical semantic segmentation structure to capture lane related information focusing on the region of interest, i.e., where lane markings are likely to exist, and also enable training on multiple heterogeneous datasets. An optimization-based lane inference method to directly estimate all parameters (i.e., lateral offsets, heading angle, and curvature) of multiple lanes in real-time. Especially, an ingeniously designed cost function considering road/lane geometric models is proposed to formulate the parameter estimation problem.

MODULES

- Subnet
- Hierarchical Semantic Segmentation
- Lane Parameter Estimation
- Lane Identification

SUBNET

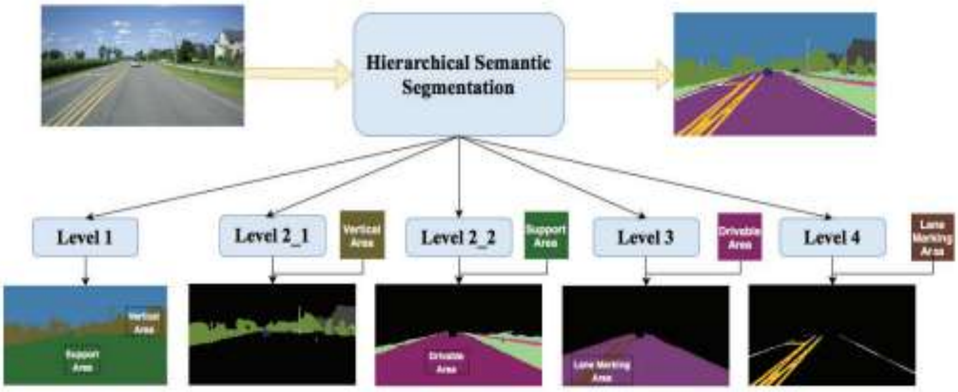
Convolution performs blurring, sharpening, edge detection, noise reduction etc., Deconvolution is a computationally intensive image processing technique that is being increasingly utilized for improving the contrast and resolution of digital images captured in the microscope. Bilinear up sampling is the process of creating a larger resolution image where every sample is created from bilinear filtering of a smaller resolution image.

Hierarchical Semantic Segmentation

To understand the holistic scene for reliable lane detection, semantic segmentation is selected as the base network for the street scene understanding. To capture the special semantic relationship between lanes/roads and other objects with multiple heterogeneous datasets. Semantic segmentation refers to the process of assigning a semantic label to each pixel of an image. For learning based semantic segmentation, annotated training data is an essential part that affects model performance. In our hierarchical semantic segmentation model, both scene labels and detailed lane marking labels are required, while no current open dataset provides annotations covering all the required labels. It enables training on multiple heterogeneous datasets. This network covers different types of objects (e.g., animals and birds) but most of them are not related to lane detection and can be ignored. Hence, it cannot use the structure directly for lane detection, instead, it optimize the hierarchy with a customized architecture inspired by human perception ability to highlight lane-related objects only and improve system efficiency. The relationships between different hierarchical levels are redefined, so that the proposed structure focuses on both general scene cues and lane classification.

The proposed hierarchical structure extracts the objects in each level and this structure tries to label an image of a street scene into coarse geometric classes that are useful for tasks such as navigation, object recognition, and general scene understanding. Following the backbone network, four-level classifiers are designed: in the first level, an image is divided into sky, vertical and support areas; in the second level, vertical/support areas are further divided into subclasses for scene description; in the third level, drivable areas are further classified; in the last level, detailed lane marking types are extracted from the lane marking areas.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 66-73 www.ijaem.net ISSN: 2395-5252



LANE PARAMETER ESTIMATION

Once scene labels are predicted by the hierarchical semantic segmentation, this module can use lane related labels to estimate lane parameters in the vehicle coordinate system. It assume the lane lines are largely parallel polynomials, and the lane parameters can be divided into two parts, shared/global parameters (i.e., heading angle and curvature) and unique/local parameters (i.e., offset of each lane). Other lane marking attributes such as type and color are detected separately. The multi-lane parameters of road slope is proposed for better accuracy.

The lane parameter estimation is considered as an optimization problem. The cost function is designed to reflect the physical properties of lane lines. To measure the fitting performance, one special cost function is designed considering the physics knowledge of road/lane. The proposed cost function includes two parts, called road-fitting and lane fitting, indicating how good the parameters fit the geometric profiles of road and lane, respectively.

LANE IDENTIFICATION

The Nelder-Mead Simplex (NMS) algorithm is used for lane detection. The concept of simplex, is adopted in NMS, that is, a special polytope of n+1 vertices in n dimensions. NMS maintains a simplex, and extrapolates the behavior of the cost function to find a new test point to replace the worst point in a simplex through deterministic rules, i.e., reflection, expansion, contraction, and shrink. The above simplex evolution procedure will repeat until convergence. Usually, a proper initial guess helps to avoid trapped at the local minima and speed up the searching process. Sequential information may provide a good initial parameter, which means the optimal solution obtained from the previous images can be used as the initial value for the next image. The curvature information of the road ahead the vehicle from map or navigation software can also be used. If no prior information is available, a proper range is selected for each parameter and the initial values are set to zero.

V. CONCLUSION

The proposed system is a novel lane detection algorithm with two unique ideas: 1) it predicts lane related labels from the holistic scene understanding; 2) it estimates multi-lane parameters and compensate for road slope simultaneously under an optimization framework. The advantages of both learning-based and physics-based techniques are leveraged. The proposed algorithm is trained on heterogeneous datasets and then tested with input video. The proposed method was found to achieve similar or better performance, and is more robust. In future the proposed system can be implemented with the accurate estimation of unparalleled lanes, such as lane merge and split conditions, additional operations/strategies, e.g., extra local correction or integration with map prior.

REFERENCES

[1]. M. Bertozzi and A. Broggi, "GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection," IEEE Trans. on Image Processing, vol. 7, no. 1, pp. 62–81, Jan. 1998.

[2]. M. Aly, "Real time detection of lane markers in urban streets," 2008 IEEE Intelligent Vehicles Symposium, pp. 7-12, Jun. 2008.

[3]. Y. Hou, M. Zheng, C. Liu, and C. Chen, "Learning lightweight lane detection cnns by self-attention distillation," In Proc. ICCV, Seoul, Korea, Oct.27-Nov.2, 2019, pp. 1013-1021.

[4]. M. Ghafoorian, C. Nugteren, N. Baka, et al., "EL-GAN: embedding loss driven

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 66-73 www.ijaem.net ISSN: 2395-5252

generative adversarial networks for lane detection," in ECCV. Munich, German, Sep. 8-14, 2018, pp. 256-272.

[5]. B. He, R. Ai, Y. Yan, and X. Lang, "Accurate and robust lane detection based on Dual-View Convolutional Neutral Network," 2016 IEEE Intelligent Vehicles Symposium (IV), pp.1041-1046, Jun. 2016.

[6]. S. Xu and H. Peng, "Safeguard Protected Preview Lane Keeping Control for Automated Vehicles," in AVEC, Beijing, China, Jul. 16-20, 2018, pp. 1-8.

[7]. M. Nieto, L. Salgado, F. Jaureguizar, and J. Cabrera, "Stabilization of Inverse Perspective Mapping Images based on Robust Vanishing Point Estimation," 2007 IEEE Intelligent Vehicles Symposium, pp. 315-320, Jun. 2007

[8]. H. Kong, J. Y. Audibert, and J. Ponce, "Vanishing point detection for road detection," In CVPR, Miami, FL, USA, Jun. 20-25, 2009, pp. 96- 103.

[9]. P. Meletis and G. Dubbelman, "Training of Convolutional Networks on Multiple Heterogeneous Datasets for Street Scene Semantic Segmentation," IEEE Intelligent Vehicles Symposium (IV), pp. 1045- 1050, Jun. 2018.

[10]. Nieto, L. Salgado, F. Jaureguizar, and J. Cabrera, "Stabilization of Inverse Perspective Mapping Images based on Robust Vanishing Point Estimation," 2007 IEEE Intelligent Vehicles Symposium, pp. 315-320, Jun. 2007.