The past and present of real-time lane detection method in autonomous vehicle system

Lin He*

Jiangsu University, China

*Corresponding author: rlin1he@gmail.com

Keywords: Autonomous car, lane detection, artificial intelligence, methods.

Abstract: Smart cars have been one of the most heated debated issues in the automotive industry for decades. Although various efforts have been made to make cars smart and have the ability to drive themselves, the intention to find the optimal solution has never stopped. In the process, autonomous cars are becoming more and more stable and intelligent, moving towards the ideal of a smart car. This article describes the technologies used in self-driving cars of lane detection and how they have been improved over the last few decades. Two main sections are introduced in this article, traditional methods and deep learning based methods of lane detection.

1. Introduction

Since the concept of smart driving was introduced, various sensors have been used to control the forward movement of vehicles. Such as speed sensors, displacement sensors, gyroscopes, radar, rangefinders, etc. Human vision is the most important perceptual tool people have. Reflections from external objects enter our eyes and vision cells allow these rays to be imaged in our minds. The camera was created to mimic the optics of the eye, collecting data in the form of individual pixels, which are then rearranged in some sort of sequence to form an image. This principle allows images to be processed and used as a form of digital information.

Lane detection is the task of detecting lanes on a road from a camera, helping the car to move within the specified range. [1]The trend corresponding to the road cars captured by the camera is obtained by analysing each image frame. The camera has an override effect, it captures the image in front of the car's travel position.

2. Traditional methods

Traditional lane detection methods include the fitting method [2], the Hough transform method [3], and the inverse perspective change method [4].

The binarised image is often used as the feature extraction image in traditional lane detection methods. Binarisation means taking the best threshold to divide the image into two parts, object and background, based on the grey scale features of the image when the divisibility of the two parts is maximised.

The flow of traditional methods for lane detection [5] [6] [7] is summarised below. This can be divided into 4 steps, first we pre-processing the image, next we searching for lanes, then extracting road edges, finally it will output by controller.

2.1 Pre-processing of the image

This step is the basis for the subsequent operations. In this step there are usually two sub-steps, the masking of the image and the binarisation of the image. The images captured by the autonomous vehicle camera contain a lot of information about the road outside the lane, if it is only used for lane detection, we can hide this unimportant information and only use the lane information that is valid for us, in order to reduce the interference of other information to the lane detection. Therefore the part that can be used for lane detection can be measured manually by masking and setting the rest all to black. There are many methods of binarisation of images and the Otsu method [8] used here is just one of the typical algorithms. The Otsu algorithm is a binarisation algorithm in which all pixels in an image can be divided into two parts that are either black or white, depending on whether a pixel point is greater or less than a threshold value.

2.2 Searching for lanes

Due to the constantly changing road conditions and in order to reduce the interference of multiple lanes and passing vehicles, the search for road boundaries often starts at the centre of the image and moves from the middle to the sides, searching for boundaries from the bottom to the top [9]. On contrary, Hough Transform is a popular technique to detect any shape in a global image, if you can represent that shape in mathematical form. It can detect the shape even if it is broken or distorted a little bit.

2.3 Extracting road edges

In the binarised image, the lane guide lines are separated from the road. At the edges where they meet there is a row of lines, and each point that forms this line is called a jump point. They have a black background on one side and the lane lines on the other. The lanes are generally divided into solid and dashed lines, and lines scanned directly against the dashed lines are often incomplete. Therefore, image restoration algorithms [10] for repairing the boundary lines will be used to draw a complete line based on these line segments. This will make the extracted line segments clearer and allow for more accurate modelling input later on.

2.4 To make an input-output relationship between the boundary lines and the steering control

Construct a model to establish the mathematical relationship between the repaired boundary line and the steering output. This model can be a functional model or a neural network. More than one model may be constructed in this step, as in addition to the model of the output of the image processing results. The above steps allow the information in the image to eventually be able to control the direction of the vehicle.

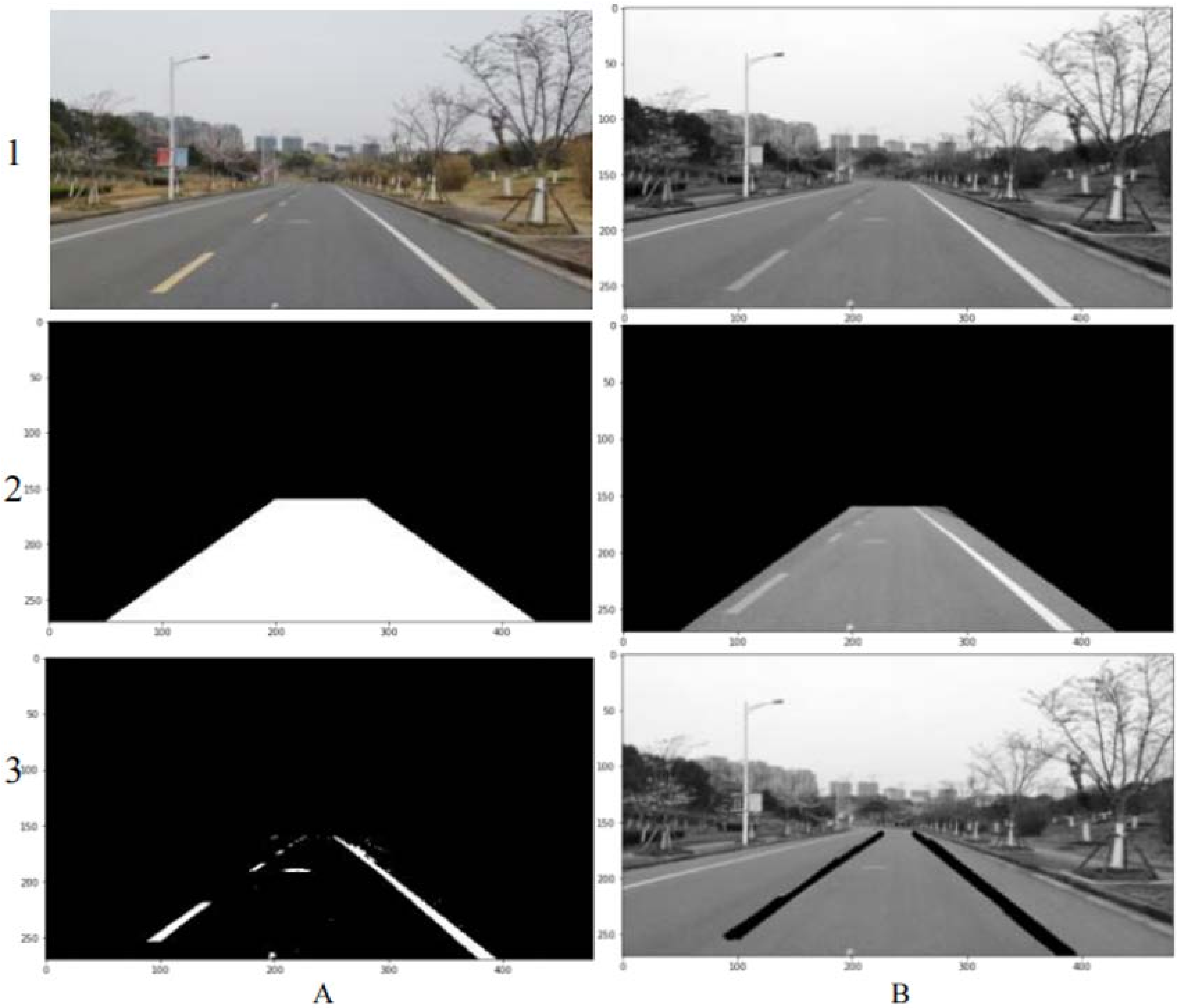

Fig 1. [11]. 1A.original photo, 1B. Grey-scale image, 2A. The masking area, 2B. Image for analyse the lane, 3A. Binarised image of 2B, 3B. Lanes (black) which are detected.

2.5 Disadvantages of traditional method

However, traditional methods mentioned above also pose a number of problems.

2.5.1 Loss of image information

The image binarisation process loses a lot of information compared to a greyscale image. This is because for each pixel there are only two possibilities (0 or 1) in a binarised image, whereas there are 255 possibilities (0-255) for each pixel in a greyscale image. The dimension of information in a greyscale image is much greater than in a binarised image, reflecting in more detail the contours of the objects in the image: a binarised image, on the other hand, will produce noise after processing due to uneven lighting, which can be filtered out using the Gaussian filter method [12].

2.5.2 Reliance on manual human work

Every traditional algorithmic model is inseparable from manual work. First, the image captured by the camera lens has to be processed, then the valid information is extracted, various algorithms are designed to build the model by repeated experiments, and finally the labelled image is collated and summarised. The whole process is completed under human operation, which requires manual adjustment of a large number of parameters increasing human workload and poor robustness.

2.5.3 Limited application scenarios

The actual road is far more complex than the experimental road, and autonomous driving has to consider not just how to travel at a uniform speed in a straight lane, but to be able to cope with various road conditions as well as people. Common roads can be any combination of the following categories: straight, curved, roundabout, intersection, zebra crossing, single lane, and dual lane, multi-lane, flat or slope, etc. We generally need to model each of these road conditions in order to make it work for the entire driving process. Each common algorithm has its own drawbacks, such as the Hough method is only applicable to straight lines, the fitting method is prone to bias, and the inverse perspective method requires a high level of input images.

Therefore, when it comes to actually applications, there are quite some limitations with of traditional methods.

3. Deep learning-based methods

With the rapid development of computer technology, high-intensity, real-time computing using large data sets has become possible. In recent years deep learning neural networks have also been used in the field of autonomous driving, especially for lane detection using cameras. In contrast to the purely mathematical computation of pixel points in traditional approaches, deep learning neural networks [13] tend to extract image features and abstract these lane lines. Nowadays, the commonly used models can be broadly classified into two categories, one is segmentation models such as SCNN [14] passing information through features and the other is target detection models such as VPGnet [15], using vanishing point guided network for lane. Three concepts of neural networks are introduced here, Convolutional neural networks [16], End-to-end neural networks [17] and Neural circuit policies [18].

3.1 CNNs (Convolutional Neural Networks) in lane detectionion

CNNs, a widely used tool for image processing, are considered a milestone in the advancement of neural networks and image processing. Feature extraction is an important function of CNNs.

It is able to extract the valid information we need in an image then it outputs a feature vector which is fed to the softmax unit to predict the road type. For example, it can extract and identify lane information in an image and then classify the lanes. Initially CNN networks were used in combination with traditional methods for lane type recognition, later CNN networks could also perform lane extraction from the original image.

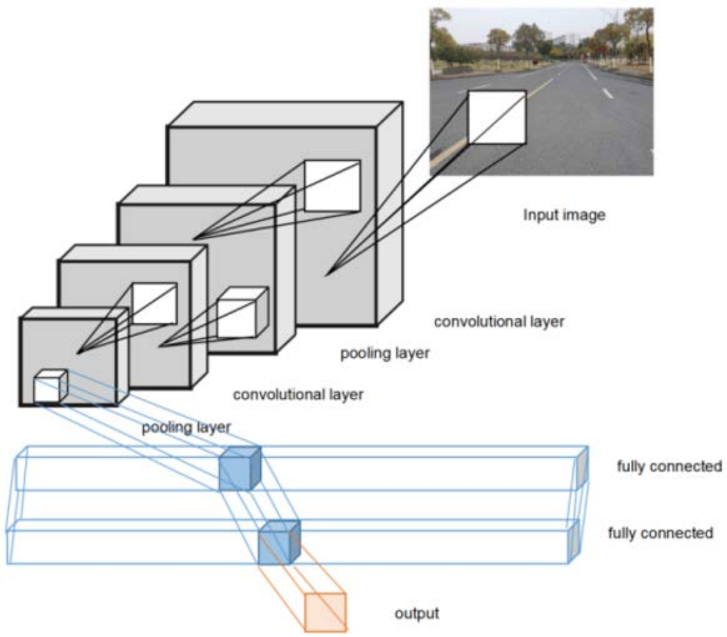

Fig 2. CNN structure of lane detection.

Here is an example of a cascaded CNN [19].

In this model consists of two neural networks in series, first CNN is trained for lane boundary instance segmentation, then extract a descriptor for each detected lane boundary and process it with second CNN. ERFNet [20] was chosen as the backbone network and the number of lane was fixed at four. In order to minimize gradient explosion and loss divergence, the curriculum learning strategy [21] has been used to achieve convergence. A descriptor is extracted for each boundary, sampling a fixed number of points from the input image, which belong to the detected lane boundary. The points extracted in this way are ordered according to the index in the original image, then arranged them into squares, which can be processed by the second neural network. As a result, this network structure did an excellent job, especially with the high frame rate of over 50fps, which is suitable for real-time road condition calculations.

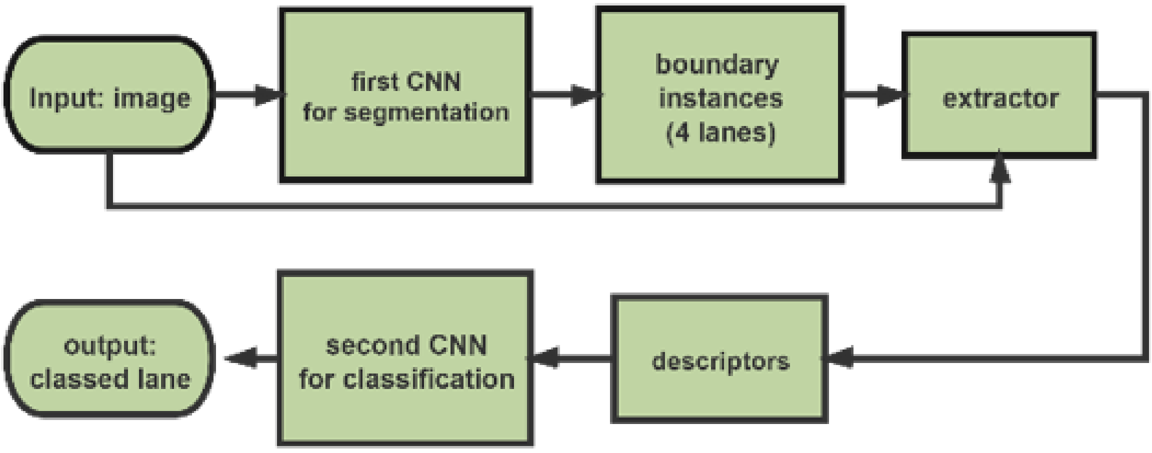

Fig 3. The structure of cascade CNN model.

3.2 End-to-end neural networks

End-to-end neural networks [22] [23] [24] are moving towards the use of fewer neurons and layers to achieve optimal control. The introduction of end-to-end learning has turned the original staged data processing into a one-step process from the input and output of the neural network. In an end-to-end neural network, the system learns the information contained in the data through a large amount of data, unconstrained by human will like a black box, and we do not fully understand the internal structure of the system or know exactly how the neurons are connected in the network structure. However, the advent of end-to-end neural networks has changed this, allowing what used to be a multi-step lane detection process to be completed in just one step. The network shown above is an end-to-end neural network whose output will be mapped to the angle of the steering engine [25] [26]. End-to-end networks are not directly related to CNNs. End-to-end networks are simply a concept of network structure, and an end-to-end network can be composed of CNNs or other neural network models.

Lanenet [27] is an end-to-end neural network widely used methods. The method designs a branched multitasking network for lane instance segmentation like [28], consisting of two parts, a lane segmentation branch and a lane embedding branch, which can be trained end-to-end. The lane segmentation branch has two output classifications, background or lane. The lane embedding branch further divides the segmented lane pixels into different lane instances and is trained using a clustering loss function that assigns a lane id to each pixel in the lane segmentation branch, while ignoring the background. This allows for multi-lane recognition.

In order to fit the curved lanes, the image is usually fitted with a bird's-eye view through an inverse perspective; the points in the ensemble of pixels output by Lanenet are not exactly on the lanes, and a better fit is obtained by using H-net to learn the parameters of the transformation matrix from the original image to transform the image.

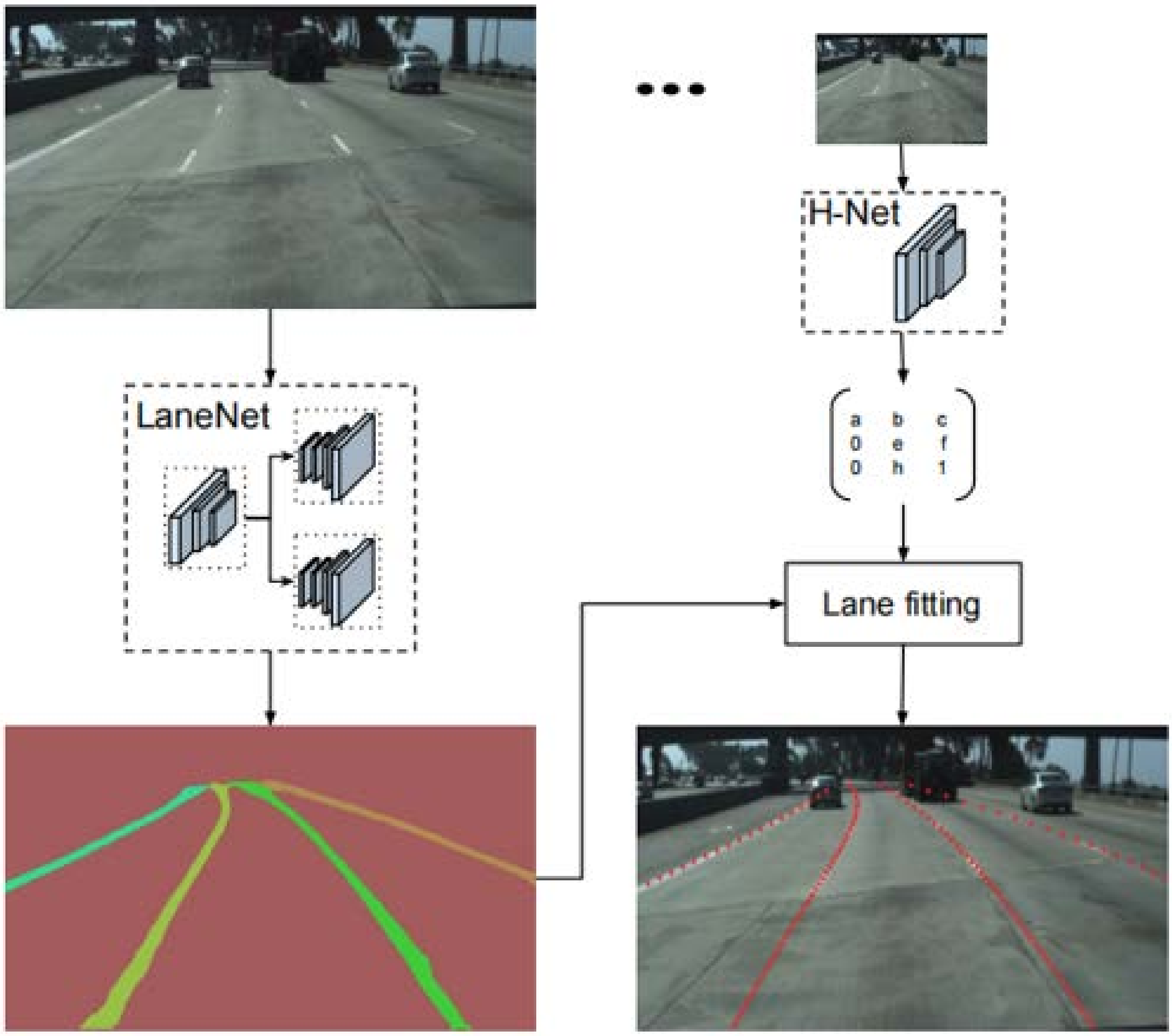

Fig 4. [26] The structure of Lanenet methods.

3.3 NCPs (Neural Circuit Policies)

Neural network models based on biological properties are gradually becoming known as a research trend. For example, ant colony algorithms and genetic algorithms have been used in practical engineering applications. Last October, researchers at MIT proposed an end-to-end neural network from biological navigation. Neural Circuit Policies (NCPs) are designed sparse recurrent neural networks based on the LTC neuron and synapse model loosely inspired by the nervous system of the organism C. elegans. [29]. Consisting of only 19 neurons in 4 layers, it is considered a revolutionary network because the network directly translates input images into output steering control in such a lightweight structure. It is worth noting that the NCP network is not directly connected to the image, but rather the road features are extracted through a layer of CNN network and used as input.

In lane detection for smart cars, the dataset is often what is recorded by the car's cameras while driving manually. Due to the richness of the dataset, end-to-end networks can be constructed from large amounts of data. The end-to-end neural network process for the entire lane detection process is roughly as follows: input image, extraction of image features, feature mapping, and output steering machine angle.

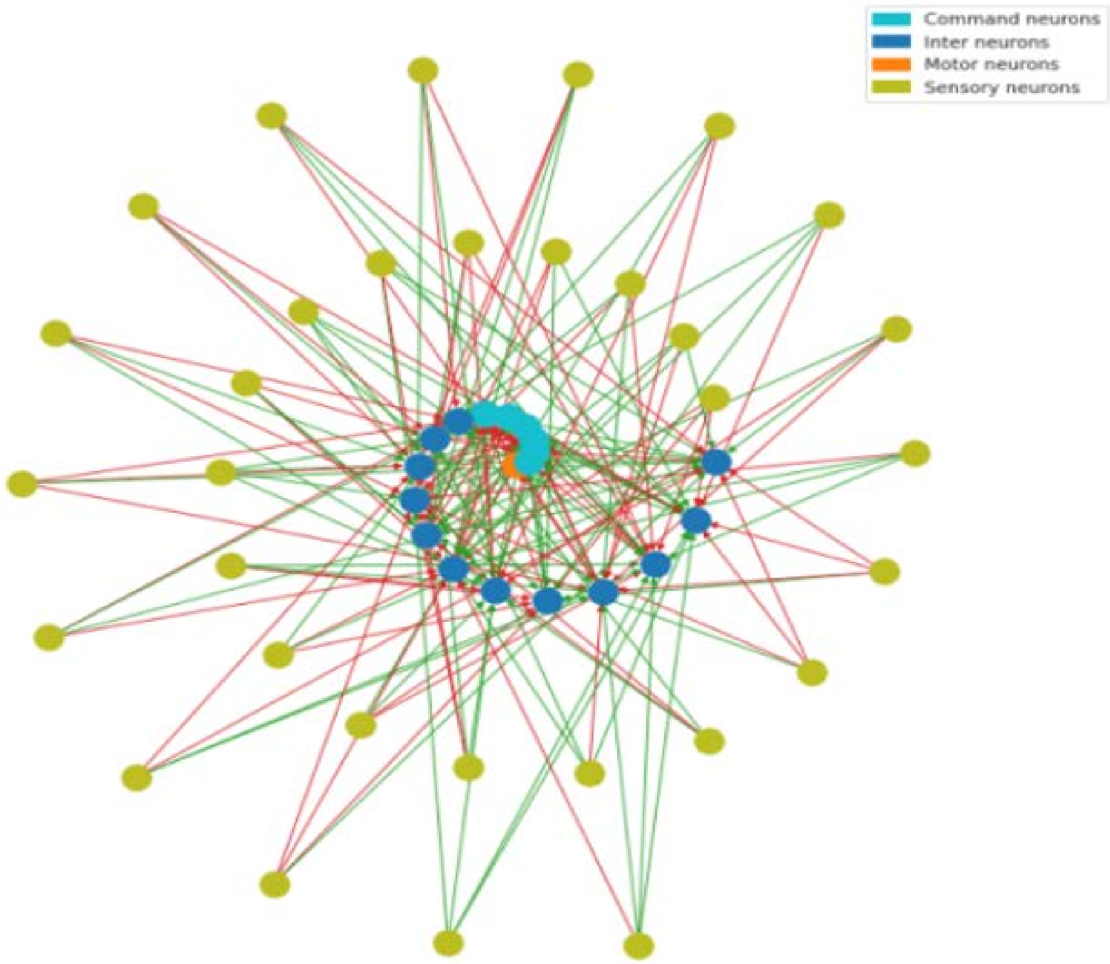

Fig 5. [30] The connection of neurons in NCPs.

The four layers of neurons are divided into perceptual neurons, intermediate neurons, command neurons and motor neurons. This corresponds to the image processing, lane detection and steering control components of autonomous driving navigation.

The main driving primitives (drive straight, left turn and right turn) have been concisely learned by an NCP's internal neural state.

3.4 Improvements and deficiencies

Despite this, deep learning still has many pressing challenges to address.

3.4.1 Multi-tasking detection

Lane detection is part of automated driving detection, in addition to detecting traffic lights, pedestrians and even other vehicles while the vehicle is in motion. Accurately detecting the same image in the same control cycle and coordinating the control of the vehicle based on the results becomes a problem.

3.4.2 Forecast of lanes

There will be discontinuities in the lane lines, which may be due to breakage, something else obscuring them. These defects will only appear for a short period of time during driving. For safety reasons the system needs to be robust and there are ways to compensate for these deficiencies and to accurately restore and predict the lane lines in advance.

3.4.3 Changing real-time road conditions

The advent of deep learning has greatly deal with the problem of limited application scenarios in traditional methods, and more effective models can be obtained with large data sets of different scenarios. In real-world automated driving, vehicles cannot always drive along a single lane in a closed loop, and there may be situations where vehicles change lanes, turn at intersections, etc. Therefore, in order to deal with these situations, we need new ways to figure them out.

4. Conclusion and future

This paper summarised methods for real-time camera lane detections. For traditional methods, we discuss the processes of how to search lane from an image, and the limitation of this old-fashioned AI. Furthermore, we also introduce deep- learning-based methods. This section includes several neural networks methods for lane detection. Each method has its own advantages and disadvantages, but there are still some problems need to be improved. Neural networks are considered to be the future of AI, because they can mimic the way of how human beings provoke their thoughts. We are expecting methods that act more like a human not a machine, which have the ability to learn and correct, detecting the lane while it can obey the traffic rules. Most importantly, it has to compact with extraordinary robustness which distinguish from increasingly tired, unfocused human drivers.

Lane detection is not a difficult task for human, but it remains to be a huge issue to function this task outside the human brain. In the past, we deem to dismantle this task and conquer them step by step is the best solution of the lane detection task. But as we looked back to the methods prompt recent years, the size of the neural network is smaller, the datasets become larger, the bond between each step is tighter and more ambiguous, which is closer to animal's nervous system. If there was a biological breakthrough in forth neural networks revolution, we are provided with opportunities to see our own "brain" detecting the lane.

5. Comparison of 3 methods in this article

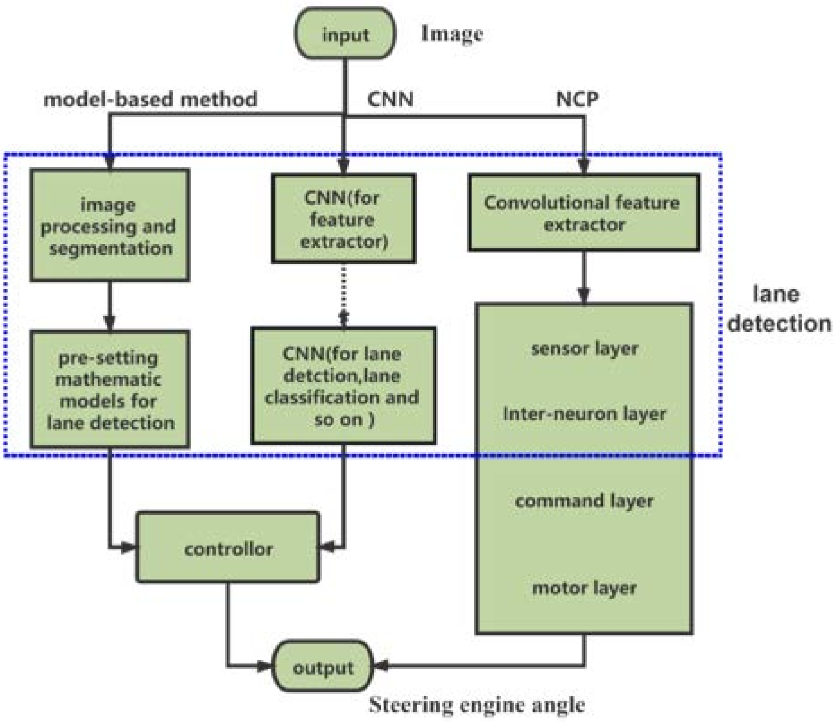

Fig 6. The comparison chart of 3 different methods mentioned for lane detection.

The flowchart above is a comparison of the methods in this paper. This flow chart of the lane detection and control process clearly shows the characteristics of each method. The traditional methods require a lot of human specification, but this is not work poorly in actual practice. The methods based on deep learning and neural networks have different training processes and outputs due to the different network constructions. The end-to-end neural network structure play a role in controller.

References

[1] https://www.paperswithcode.com/task/lane-detection

[2] Qing Li, Nanning Zheng and Hong Cheng, "An adaptive approach to lane markings detection," Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems, Shanghai, China, 2003, pp. 510-514 vol.1, doi: 10.1109/ITSC.2003.1252005.

[3] B. Yu and A. K. Jain, "Lane boundary detection using a multiresolution Hough transform," Proceedings of International Conference on Image Processing, Santa Barbara, CA, USA, 1997, pp. 748-751 vol.2, doi: 10.1109/ICIP.1997.638604.

[4] Gang Yi Jiang, Tae Young Choi, Suk Kyo Hong, Jae Wook Bae and Byung Suk Song, "Lane and obstacle detection based on fast inverse perspective mapping algorithm," Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics. 'Cybernetics evolving to systems, humans, organizations, and their complex interactions' (cat. no.0, Nashville, TN, USA, 2000, pp. 2969-2974 vol.4, doi: 10.1109/ICSMC.2000.884452.

[5] Risack, R., Mohler, N., & Enkelmann, W. (n.d.). A video-based lane keeping assistant. Proceedings of the IEEE Intelligent Vehicles Symposium 2000 (Cat. No.00TH8511).

[6] H. Nishihara, A. Kojima, H. Murakoshi and S. Ishijima, "The multi-layer neural network applied to a car detection system," [1992] Proceedings IEEE International Workshop on Robot and Human Communication, Tokyo, Japan, 1992, pp. 88-92, doi: 10.1109/ROMAN.1992.253919.

[7] D. M. Gavrila, U. Franke, C. Wohler and S. Gorzig, "Real time vision for intelligent vehicles," in IEEE Instrumentation & Measurement Magazine, vol. 4, no. 2, pp. 22-27, June 2001, doi: 10.1109/5289.930982.

[8] Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, 9(1), 62–66. doi:10.1109/tsmc.1979.4310076

[9] Y. Hamada, T. Nakamori, N. Ishikawa and M. Nakajima, "Development of in-vehicle system for evaluating the "quality of driving"," IVEC2001. Proceedings of the IEEE International Vehicle Electronics Conference 2001. IVEC 2001 (Cat. No.01EX522), Tottori, Japan, 2001, pp. 27-30, doi: 10.1109/IVEC.2001.961721.

[10] C. Kreucher and S. Lakshmanan, "LANA: a lane extraction algorithm that uses frequency domain features," in IEEE Transactions on Robotics and Automation, vol. 15, no. 2, pp. 343-350, April 1999, doi: 10.1109/70.760356.

[11] https://www.analyticsvidhya.com/blog/2020/05/tutorial-real-time-lane-detection-opencv/

[12] G. Deng and L. W. Cahill, "An adaptive Gaussian filter for noise reduction and edge detection," 1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference, San Francisco, CA, USA, 1993, pp. 1615-1619 vol.3, doi: 10.1109/NSSMIC.1993.373563.

[13] Tang, J., Li, S., & Liu, P. (2020). A Review of Lane Detection Methods based on Deep Learning. Pattern Recognition, 107623. doi:10.1016/j.patcog.2020.107623

[14] ArXiv: 1712.06080v1 [cs.CV]

[15] Lee, S., Kim, J., Yoon, J. S., Shin, S., Bailo, O., Kim, N.… Kweon, I. S. (2017). VPGNet: Vanishing Point Guided Network for Lane and Road Marking Detection and Recognition. 2017 IEEE International Conference on Computer Vision (ICCV). doi:10.1109/iccv.2017.215

[16] Zeiler, M. D. & Fergus, R. Visualizing and understanding convolutional networks. In European Conference on Computer Vision 818–833 (2014).

[17] Lecun, Y., Cosatto, E., Ben, J., Muller, U. & Flepp, B. Dave: Autonomous Of-road Vehicle Control Using End-to-end Learning Technical Report DARPA-IPTO Final Report (Courant Institute/CBLL, 2004); https://cs.nyu.edu/~yann/research/dave/

[18] Lechner, M., Hasani, R., Amini, A., Henzinger, T. A., Rus, D., & Grosu, R. (2020). *Neural* circuit policies enabling auditable autonomy. Nature Machine Intelligence, 2(10), 642–652. Doi: 10.1038/s42256-020-00237-3

[19] ArXiv: 1907.01294v2 [cs.CV]

[20] Romera, E., Alvarez, J.M., Bergasa, L.M., Arroyo, and R: Erfnet: Efficient residualfactorized convnet for real-time semantic segmentation. IEEE Transactions on ITS

[21] Bengio, Y., Louradour, J., Collobert, R., Weston, J.: Curriculum learning. In: Pro-ceedings of the 26th annual international conference on machine learning. (2009)

[22] Xu, H., GAO, Y., Yu, F. & Darrell, T. End-to-end learning of driving models from large-scale video datasets. In Proc. IEEE Conference on Computer Vision and Pattern Recognition 2174–2182 (2017).

[23] Bojarski, M. et al. End to end learning for self-driving cars. Preprint at http://arXiv.org/abs/1604.07316 (2016).

[24] https://medium.com/@zsyed350/end-to-end-learning-for-self-driving-cars-made-simple-9defc04d360e

[25] Amini, A., Paull, L., Balch, T., Karaman, S. & Rus, D. learning steering bounds for parallel autonomous systems. In IEEE International Conference on Robotics and Automation (ICRA) 1–8 (2018).

[26] Neusser, S., Nijhuis, J., & Spaanenburg, L. (n.d.). Developments in autonomous vehicle navigation. CompEuro 1992 Proceedings Computer Systems and Software Engineering.

[27] ArXiv: 1802.05591v1 [cs.CV]

[28] D. Neven, B. De Brabandere, S. Georgoulis, M. Proesmans, L.Van Gool, Fast Scene Understanding for Autonomous Driving. Deep Learning for Vehicle Perception, workshop at the IEEE Symposiumon Intelligent Vehicles, 2017.

[29] Milford, M. (2020). C. Elegans inspires self-driving cars. Nature Machine Intelligence.

[30] https://colab.research.google.com/drive/1-mZunxqVkfZVBXNPG0kTSKUNQUSdZiBI?usp=sharing