Visual Perception Deep Drive Model for Self-Driving Car

Waleed Razzaq*, Usman Arif, and Zia Mohi U Din

***Abstract*— Self driving cars are the need of future technology, there are many companies that are trying to perfect this particular project but there are still some deficiencies there. Most of the companies are using Expensive sensors like RADAR and LiDAR to get the idea of environment, which are very hard to use and need a lot of processing power. Our project focuses on using only visual aid to drive a car, particularly following the lane of the road. We trained a model using Convolutional Neural Network (CNN), in a simulated environment and tested the model in the same environment.**

***Index Terms*— Convolutional Neural Network, Simulated Environment, RADAR, LiDAR.**

I. Introduction

CURRENTLY Advanced driver assistance systems reinforce driver to take decisive action or trigger certain operation of vehicle to avoid injury or damage [1]. These systems put base for autonomous vehicle's era from one autonomous action to fully autonomous. Human errors are enormous reason for road accidents daily happening [2]. Scientists and researchers are working for a long time on autonomous vehicles to decline this causality. Many contemporary techniques are developed in both cyber and physical fields to increase the efficiency and protection of vehicles. Cyber systems include recognition, networking and communicating while physical systems cover sensing environment and actuating [3]. Due to continuous variability of environment, it is necessary for vehicles to be robust, efficient and precautious to handle the problems generated instantly [4]. The destination point, vehicle's active position and the path to be followed is determined by using GIS (Geographic information system). However, the path to be followed is not enough for vehicles to arrive at destination. Vehicles have to look after the variable environment, data available, update current position data and make decisions accordingly. The responsive ability is important in terms of security and accuracy. Lane keeping of vehicles is figured as elemental response to execute GIS information. Lane keeping system is necessary for vehicle to proceed at road. Vision based systems for lanes include inverse perception mapping (IPM), path spotting with edge based and detection by colour.

R. Marino et.al [5] used proportional integral derivative (PID) controller to detect the lanes at road and perform lane following by autonomous vehicle in undetermined environment. Vision based driver assistant systems has been proposed to monitor the behaviour of driver and assist him in driving precisely [6]. Labayrade et. al worked on generating the commands when vehicle try to departure lanes [7] while [8] proposed technique for centring the car between lanes. C.Y Kuo et.al presented lane detection and vehicle control algorithm for autonomous vehicle using radar sensors and image sensors. These all models use RADAR or LiDAR to fulfil the tasks.

With innovations in technology autonomous vehicles (AVs) have also developed. AVs are presumed to replace human drivers in coming decades with intelligent systems, which will be more secure, comfortable and efficient in terms of vision operation, perceiving environment and making decisions [10]. Machine learning (ML) can be used for in developing autonomous model. ML trained models cannot only detect the lanes and categorize road indicators [11] but in some situations ML techniques are useful to control the motion of systems and end to end neural networking techniques is being used due to empirical gain and resilience [12]. In a recent research T. Onishi et.al [13] proposed autonomous driving system with end-to-end learning method by using path following function. Data was collected from the track and generated deriving conduct using convolutional neural network (CNN) trained by data to follow lane. Many researchers used neural network for developing AVs [14-15]. In this paper, we have trained our neural network using audacity simulator and checked our results in the same simulator. We have taken a 120*120-resolution photo from the screen of a simulator then we convolved this photo five times and used pooling to bind the convolved structure. Then we ran our retrained neural network, which was trained, on the same simulator. Results are tested on the same simulator, and as accuracy/lose graph is shown in result section.

II. Methodology

The focus of this paper is to design a model for Car to control it between the lanes of road. Definition of state space(S), action space (A) are prerequisites for the model to train.

A. State space (S)

State space is the observed environment in which a model to be trained, is going to perform its actions. Many sensors have been designed to perform the task of observing state space these includes LIDAR, IMUs, IR depth sensor and GPS location sensors. These sensors are very expensive, and are not easy to use.

Electrical Engineering University of Gujrat, Gujrat, Pakistan (correspondence e-mail: waleedrazzaq72011@gmail.com).

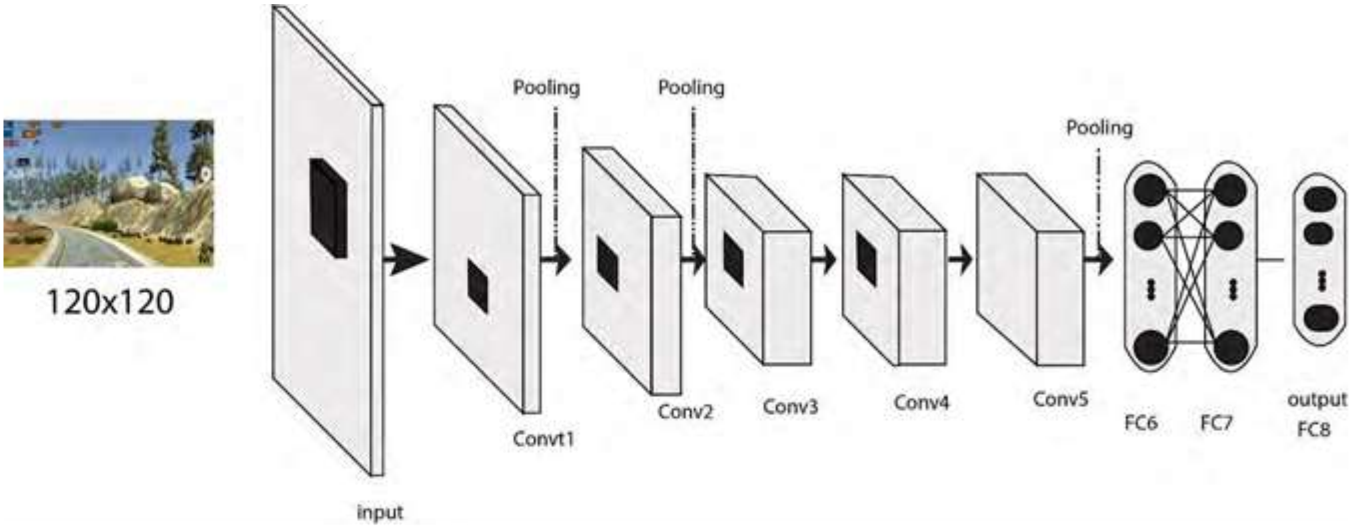

Fig. 1. Neural network structure of our algorithm

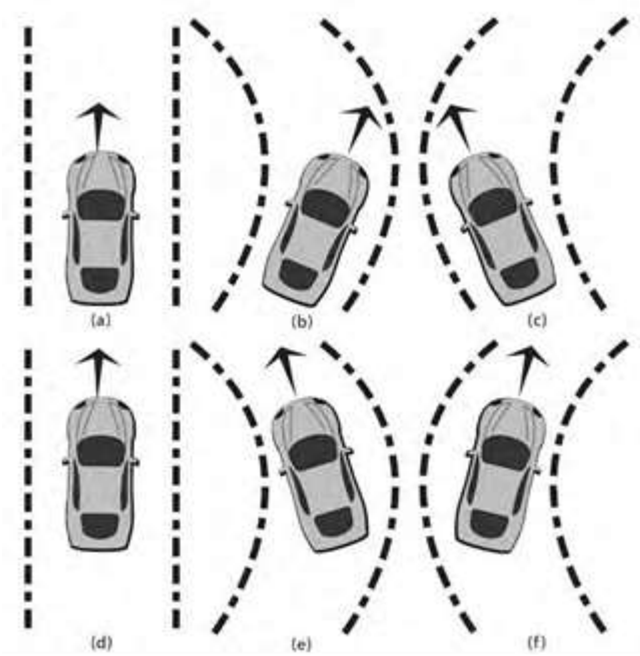

Fig. 2. Illustrations of systems of lane following (a) Vehicle going straight (b) turning toward left (c) turning toward left (d) (e) (f).

In this paper, we shall focus on monocular camera input instead of above-mentioned sensor, which is very cost effective, and produce nearly identical results. Combining the camera with car to control the speed and steering is sufficient to observe environment like humans. There are many data set available for Car to train but we have recorded the data set directly from our Simulator for better optimization. Data has been used in compressed form as algorithm is so designed to extract only the rich information (see Fig.1 and Fig. 2).

B. *Action space (A)*

The Key focus of this paper is to detect the lanes in road and find the middle point lane of those lanes. To move the car between these lanes three dimensions of action space has been defined for car [left, forward, right].

C. *Model Architecture*

In order to train our model, we have used AlexNet CNN. It has 60 million parameters and 650,000 neurons. AlexNet consist of 5 convolutional layers and 3 fully connected layers. The first two convolutional layers are followed by max pooling. The third and fourth layers are directly connected to each other. Another max-pooling layer follows the fifth. The internal structure of AlexNet is shown in Fig. 3.

D. *Preprocessing*

Working directly with recorded frames, which are 1280x720 pixel images with a 64-colour palette, can be challenging in terms of computation and memory requirements. We apply a basic preprocessing step aimed to reducing the input dimensionally. We change the resolution of image to 120x120 and into gray scale in order to reduce its demand of computation and more important reducing its dimensions. This was necessary to remove flickering that is present while other objects are present in old frames.

E. *Model Training*

To train the model we have used Intel Xeon e3 1226 v3 processor with 16 GB ram and GTX 980 G1 4G GPU. At learning rate of 1e-3, System takes almost 13 hours to train the model with 96.75 percent accuracy and 0.082 percent Loss rate has been achieved. The Accuracy / Loss graph of model is given in results section.

F. *Model Testing*

To test the model same system has been used. Continuously capturing image from simulator is subjected to our trained model give us three predictions in which the biggest prediction is calculated and performed the relative action on the car. Model training and testing chart has been showed in Fig. 3.

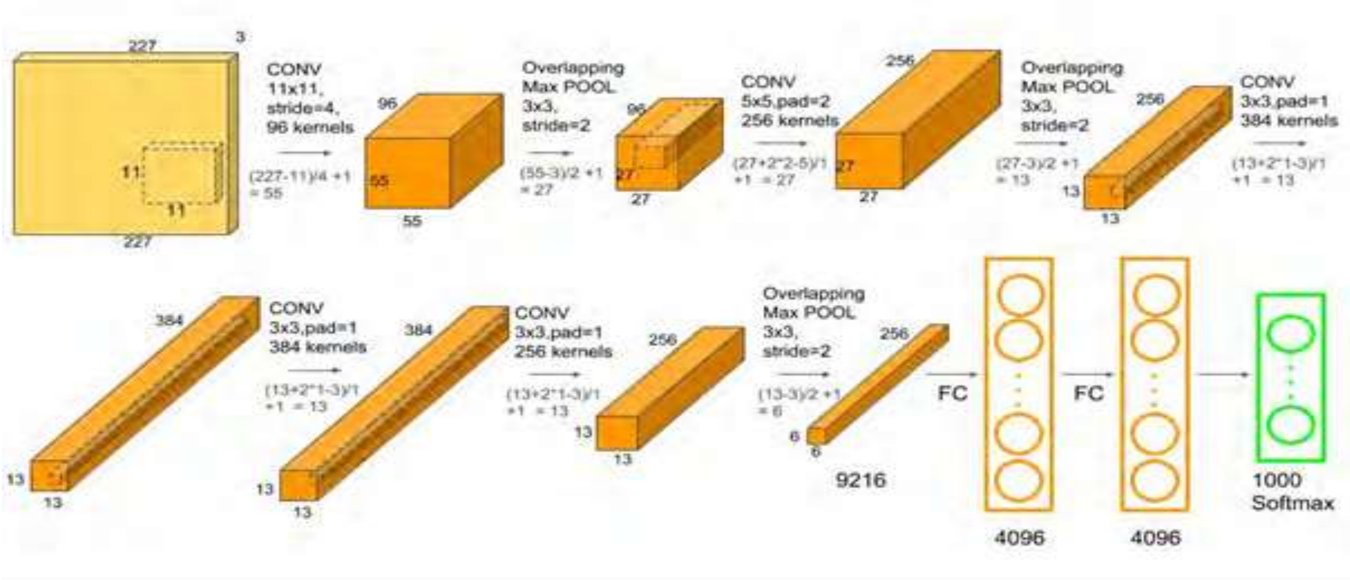

Fig. 3. Internal structure of the AlexNet CNN shown with hidden processes and calculations.

III. RESULTS

Model has been tested several times in simulator with different scale of resolution and results are mostly identical.

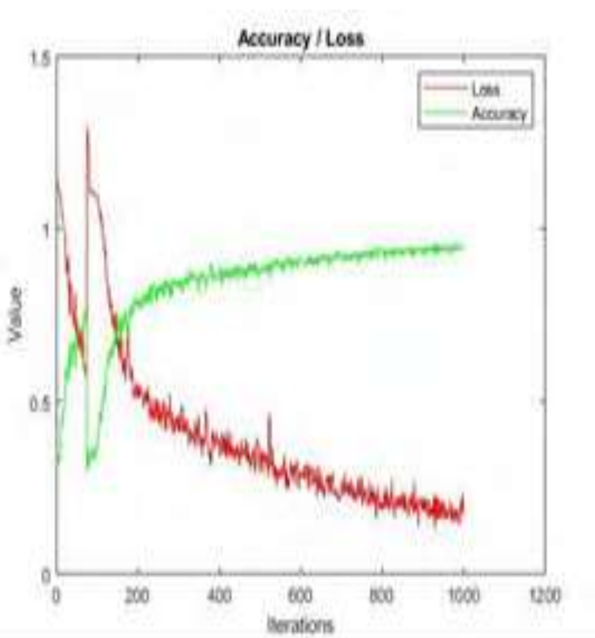

Fig. 4. Accuracy over loss graph of our model.

The model successfully detected two lines and fined the middle line between them as shown in figure. The Car follows the line that is found by our model. If the car is on the line the forward action is predicted by model and performed by the car. When the car go off the line the model predicts the action and car takes an action accordingly (see Fig. 4). These actions with visualization are shown in Fig. 5.

IV. CONCLUSION

Presented model was capable of predicting the next action of the car to follow the lane on the road and was able to drive the car in the simulator so that it follows the predicted middle lane. The whole model was trained using AlexNet CNN on a simulated environment. And was tested on the same simulator.

REFERENCES

[1] Bar Hillel, A., Lerner, R., Levi, D., & Raz, G. Recent progress in road and lane detection: a survey. Machine Vision and Applications, 25(3), 727–745, 2012.

[2] Global Status Report on Road Safety 2015; World Health Organization, Geneva 2015

[3] X. Hu, H.Wang, and X. Tang, “Cyber-physical control for energy- saving vehicle following with connectivity,” IEEE Trans. Ind. Electron., vol. 64, no. 11, pp. 8578–8587, Nov. 2017.

[4] C. Hu, R. Wang, and F. Yan, “Integral sliding mode-based composite non-linear feedback control for path following of four- wheel independently actuated autonomous vehicles,” IEEE Trans. Transport. Electrific., vol. 2, no. 2, pp. 221–230, Jun. 2016.

[5] Marino, R., Scalzi, S., & Netto, M. “Nested PID steering control for lane keeping in autonomous vehicles. Control Engineering Practice” 2011.

[6] McCall, J. C., & Trivedi, M. M. Video-Based Lane Estimation and Tracking for Driver Assistance: Survey, System, and Evaluation. IEEE Transactions on Intelligent Transportation Systems, 7(1), 20–37, 2006.

[7] Labayrade, R., Douret, J., Laneurit, J., Chapuis, R.: A reliable and robust lane detection system based on the parallel use of three algorithms for driving safety assistance. IEICE Trans. Inf. Syst (2006)

[8] Wu, S., Chiang, H., Perng, J., Chen, C., Wu, B., Lee, T.: The heterogeneous systems integration design and implementation for lane keeping on a vehicle. IEEE Trans. Intell. Transp. Syst. (2008)

[9] C.Y. Kuo, Y.R. Lu and S.M. Yang, “On the Image Sensor Processing for Lane Detection and Control in Vehicle Lane Keeping Systems”, MDPI Sensors 2019

[10] D. Lenz, T. Kessler, and A. Knoll. "Stochastic model predictive controller with chance constraints for comfortable and safe driving behaviour of autonomous vehicles." Intelligent Vehicles Symposium 2015.

[11] Mogelmose, M. M. Trivedi, T. B. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey", IEEE Transactions on Intelligent Transportation Systems, 13(4):1484-1497, 2012

[12] Simmons, B., Adwani, P., Pham, H., Alhuthaifi, Y., & Wolek, A. Training a Remote-Control Car to Autonomously Lane- Follow using End-to-End Neural Networks. 2019 53rd Annual Conference on Information Sciences and Systems (CISS), 2019.

[13] Onishi, T., Motoyoshi, T., Suga, Y., Mori, H., & Ogata, T. End-to-end Learning Method for Self-Driving Cars with Trajectory Recovery Using a Path-following Function. 2019 International Joint Conference on Neural Networks (IJCNN), 2019.

[14] H. Xu, Y. Gao, F. Yu, and T. Darrell. End-to-end learning of driving models from large-scale video datasets, 2016.

[15] F. Codevilla, M. Muller, A. Dosovitskiy, A. Lopez, and V.Koltun. End-to-end driving via conditional imitation learning, 2017.

Fig. 5. Actual Results of Simulator showing middle predicted line and car following that line