A Novel Autonomous Taxi Model for Smart Cities

NS Rajput, Mukesh, A Mishra, A Sisodia
Department of Electronics Engineering
Indian Institute of Technology (Banaras Hindu University)
Varanasi, India
nsrajput.ece@iitbhu.ac.in

I Makarov
School of Data Analysis and Artificial Intelligence,
National Research University Higher School of Economics,
Moscow, Russia
iamakarov@hse.ru

***Abstract*—Autonomous taxies are in high demand for smart city scenario. Such taxies have a well specified path to travel. Therefore, these vehicles only required two important parameters. One is detection parameter and other is control parameter. Further, detection parameters require turn detection and obstacle detection. The control parameters contain steering control and speed control. In this paper a novel autonomous taxi model has been proposed for smart city scenario. Deep learning has been used to model the human driver capabilities for the autonomous taxi. A hierarchical Deep Neural Network (DNN) architecture has been utilized to train various driving aspects. In first level, the proposed DNN architecture classifies the straight and turning of road. A parallel DNN is used to detect obstacle at level one. In second level, the DNN discriminates the turning i.e. left or right for steering and speed controls. Two multi layered DNNs have been used on Nvidia Tesla K 40 GPU based system with Core i-7 processor. The mean squared error (MSE) for the detection parameters viz. speed and steering angle were 0.018 and 0.0248 percent, respectively, with 15 milli seconds of real-time response delay.**

***Index Terms*—Autonomous Taxi, Deep Learning, Driver assistance systems, Lane detection, Smart City formatting.**

I. Introduction

In the early 1920s the first attempt was made towards driverless vehicles [1] and got momentum in the 1980s when researchers managed to develop automated highway systems [2]. After this many work towards semi-autonomous and autonomous vehicles were made largely in Germany and U.S. The U.S. defense agency, Defense Advanced Research Projects Agency (DARPA), which is responsible for the development of emerging technologies organized driverless car competition many times which also encourage the research in the direction of autonomous vehicle navigation. Renowned car manufacturers like, Bavarian Motor Works (BMW), Ford, General Motors (GM), Google/ Waymo, and Tesla are also trying to build autonomous cars [3].

The main perceptual cues include road color and texture, lane markings, and road boundaries. So, semi-autonomous and autonomous vehicles are relying on the same perceptual cues as humans do. Till now road and lane perception via the traditional cues considered to be most effective for autonomous driving. We have to include the extent of the road the number and positions of lanes, merging, splitting and ending lanes and roads in urban, rural and highways scenarios for understanding of road and lane. Although many work related to this has been made in recent years still it is beyond the current perceptual systems. There are varieties of sensors used for road/ lane detection such as, camera, stereo, LIDAR, etc [4]. Vehicle mounted cameras based on autonomous driving method is one of the best method as per the current research methodology. Therefore, machine learning algorithms becomes helpful and efficient method to extract the appropriate features from the video frames and detect road/ lane and control as well as steer the vehicle autonomously. Deep Neural Network (DNN) algorithm is well capable of learning these complex features and improved the pattern recognition techniques tremendously. It is also emerged as much powerful artificial intelligence tool in machine learning [5]. DNN architecture is a multilayer architecture of simpler layers stacked upon one another such that it can learn complex nonlinear input – output mappings easily. Therefore, DNN can be utilized for navigating vehicles autonomously. In 2016, Bojarski et al. proposed CNN architecture for learning the frames captured from single front facing camera images such that it can map the steering angles of the steering wheel of the car [6].

In this paper, a novel autonomous taxi model based on a hierarchical Deep Neural Network (DNN) architecture has been proposed for smart city scenario. In first level, the proposed DNN architecture classifies the straight and intersections/ turning of road. Simultaneously, a parallel DNN detects obstacle at this level. On second level, the DNN discriminate the turning viz. left or right and accordingly steering and speeds are adjusted. The travelling path information has already been given to the autonomous taxi and the GPS data is pre-fed to the vehicle. Now, the autonomous taxi has to take care of obstacles and lane keeping in onboard driving. Rest of the paper is as follows.

The motivation about the work and related background theory is given in Section II. The detailed experimental setups and the proposed framework have been described in Section III. Results have been shown and discussed in Section IV and finally the conclusion is mentioned in Section V.

II. Motivation and Background Theory

There are various vision-based lane-detection algorithms have been developed till date. They usually utilized different road models such as 2D or 3D, straight or curve etc. and different techniques such as Canny edge detector, Hough transform functions, template matching, neural networks, etc.

A. Feature Extraction Using Image Processing

The main objective of image processing is to extract information about the position of vehicle with respect to road from the captured video images. There are mainly two processes involved for the implementation of the above mention objective. They are namely the pre-processing process and the lane detection process. In the pre-processing process we generally remove noise and make the image sharper and in lane detection process we generally detect the desired lane of the vehicle in order to obtain the look-ahead distance and the lane angle. The lane detection process is based on real-time data of video sequences taken from a vehicle driving on the road. There are four processing steps involved for the lane detection algorithm namely image segmentation, edge detection, Hough Transform, and lane tracking.

B. Canny Edge Detection

This is a popular statistical method for detection of lines or edges. Brief details of the mathematical steps involved in 'Canny Edge Detection' technique are given below.

1. Firstly, 'Gaussian filter' (Eq. 1) is applied to remove noise from the image.

$$H_{ij} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(i-(k+1))^2 + (j-(k+1))^2}{2\sigma^2}\right); 1 \le i, j \le (2k+1) \qquad (1)$$

2. The intensity gradient is measured by Eq. 2 as,

$$G = \sqrt{G_x^2 + G_y^2}$$
$$\theta = a\tan 2(G_y, G_x) \qquad (2)$$

3. Now, non- maximum suspension technique is used to detect the edges, double threshold technique to determine potential edges and hysteresis technique to determine the weak and strong edges.

III. Experimental Setup

In this experiment, we have trained DNN architecture to navigate vehicles autonomously. DNN architecture which we are using in our vehicle is trained in such a way that it can decide where to navigate. Now vehicle is moving in a straight road with normal speed after few meters at the intersection point speed of the vehicle decreases which will indicates that there is a curve. Then the DNN architecture was trained to meet both of the requirements viz. steering and speed controls.

A. Dataset and Data Processing

The dataset used in this paper was prepared after recording our own videos. The dataset contains 2 hours of driving data, including 17 video clips recorded at 60 frames per second and some other measurements such as speed, lane detection, obstacles detection etc. The image frames are of size 640 x 180 pixels and are cropped from our own recorded video frames. The DNN architecture was then trained to map both of the requirements viz. steering and speed controls. Based on training data, this model is self-optimized. Thus preparing dataset is the very important task. To give a maximum overall performance, the intermediate parameters are self-adjusted to be optimal. Video of dash cam was taken while human driving. This dataset contains two important parameter frame and label. Frame i.e. raw image can be obtained from that video.

B. Proposed Framework

The proposed framework consists of hierarchical DNN architecture as shown in Fig. 1.

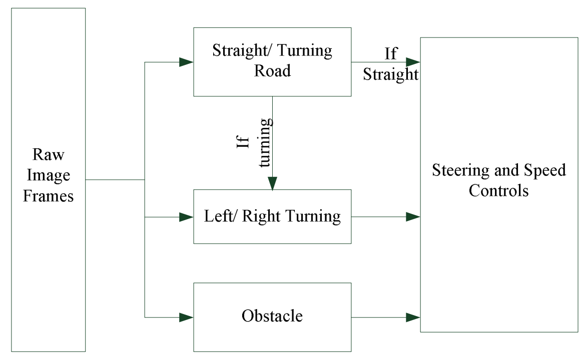
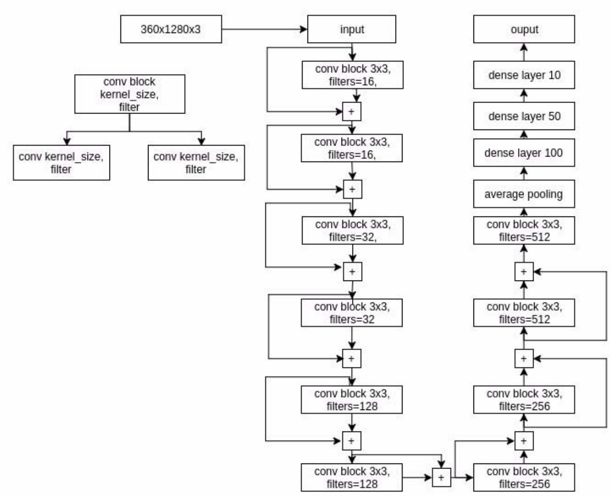


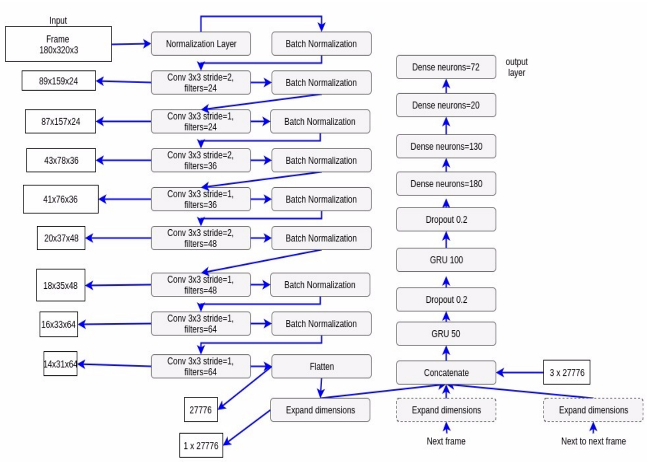
Fig. 1. Block diagram of the proposed framework.

C. Proposed DNN Architecture

DNN – I determine whether the road is straight or there is any turning/ intersection. We have used eleven convolutional layers to extract spatial features from single image frames (training input) of resolution $1280 \times 360 \times 3$, each. The detailed information is shown in Fig. 2 (a). Further, DNN – II is used for steering and speed controls. The training input for this DNN has 3 – image frames of resolution $180 \times 320 \times 3$, each. These inputs are a time-series input features of 3 steps and then passed to the double layered Gated Recurrent Unit (GRU) network. The output of the last GRU layer is then passed to a MLP classifier of 4 dense (fully-connected) layers.



(a) DNN – I



(b) DNN – II

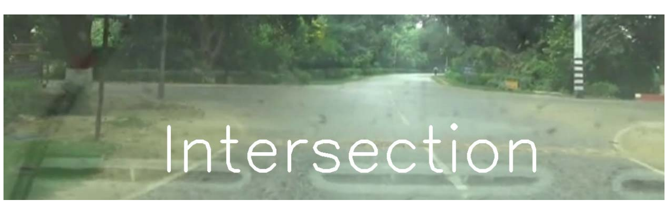
Fig. 2. Proposed architectures (a) DNN – I, (b) DNN – II.

IV. Results and Discussions

The straight and turning predictions have been shown in Fig. 3 (a) – (e) for various situations as predicted by the proposed DNN architecture. The subfigures show various conditions and their onboard predictions via our proposed DNN architectures. In Fig. 3 (a) and (b) the straight road has been predicted correctly by the proposed architecture. The 4-way intersection and 3-way intersection have been predicted accurately in Fig. 3 (c) and (d), respectively. The proposed DNN architecture results perfectly in rainy season also (Fig. 3 (e) shows the result of raining condition).

(a) Straight road predicted by proposed DNN.

(b) Straight road predicted by proposed DNN.



(c) 4-way junction predicted by proposed DNN.

(d) 3-way junction predicted by proposed DNN.

(e) 3-way Junction in raining season predicted by proposed DNN.

Fig. 3. (a) – (e) Various prediction results for straight and turning detection as obtained from our proposed DNN framework.

Steering angle results of the proposed DNN framework have been shown in Fig. 4 (a) – (c). The blue marker indicates the predicted turning. Fig. 4 (a) and (b) show the predicted steering angle for pedestrians approaching to vehicle whereas Fig. 4 (c) show the predicted steering angle for an obstacle (Car) on the way to the autonomous vehicle.

(a) Turning as predicted by the proposed DNN when pedestrian approaching to vehicle (away).

(b) Turning as predicted by the proposed DNN when pedestrian approaching to vehicle (closer).

(c) Turning as predicted by the proposed DNN when obstacle (Car) on the way to vehicle.

Fig. 4. (a) – (c) Various prediction results for steering angles as obtained from our proposed DNN framework.

Table I, tabulate various cases of turning angle prediction of the steering of autonomous taxi as predicted by the proposed DNN architecture.

TABLE I. STEERING ANGLE

Cases	Obstacles	Predicted	Actual
Fig. 4 (a)	Pedestrian approaching to vehicle (away)	85 °	87 °
Fig. 4 (b)	Pedestrian approaching to vehicle (closer)	0 °	0 °
Fig. 4 (c)	Obstacle (Car) on the way to vehicle	2 °	1 °

V. CONCLUSIONS

In smart city scenarios, autonomous taxies will be playing significant role for implementing intelligent public transportation systems. In this direction, low cost and environment friendly solutions are electrically powered autonomous taxies with single front camera based computationally less complex systems that are inspired by human driving capability. Humans take visual perceptions and draw inferences for better driving by using their past experiences and estimate the optimal driving path. In this work, we have demonstrated a novel and simpler obstacle avoidance and path planning system for implementing autonomous taxies using deep neural networks and by mimicking human decision making approach for efficient and autonomous driving. Such intelligent systems are more efficient, robust and promising.

ACKNOWLEDGMENT

The financial support, in part, received from the NCC LAB, Department of Electronics Engineering, IIT (BHU), India (Grant no. IS/ST/EC-13-14/02), from IBM, India (Grant no. R&D/IBM/SBApp/Electronics/15-16/07) and Nvidia, USA for grant of Tesla K 40 GPU under Hardware Support Grant, is gratefully acknowledged. Part of this work has been prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy by the Russian Academic Excellence Project '5-100'.

REFERENCES

[1] Weber M (2014) Where to? a history of autonomous vehicles. Available at http://www.computerhistory.org/atchm/where-to-ahistory-of-autonomous-vehicles.

[2] Fenton RE, Mayhan RJ (1991) Automated highway studies at the ohio state university-an overview. IEEE Trans Vehicular Technol 40:100–113.

[3] B. Paden, M. Cap, S. Z. Yong, D. Yershov, and E. Frazzoli, “A survey of motion planning and control techniques for selfdriving urban vehicles,” IEEE Trans. on Intel. Vehicles, vol. 1, no. 1, pp. 33–55, March 2016.

[4] Aharon Bar Hillel, Ronen Lerner, Dan Levi, Guy Raz, “Recent progress in road and lane detection: a survey,” Machine Vision and Applications (2014) 25:727–745DOI 10.1007/s00138-011-0404-2.

[5] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, May 2015.

[6] M. Bojarski et al. (2016). “End to end learning for self-driving cars.” [Online]. Available: https://arxiv.org/abs/1604.07316.