INTELLIGENT SPEED ADAPTIVE SYSTEM USING IMAGE REGRESSION METHOD FOR HIGHWAY AND URBAN ROADS

Bhavesh Sharma[1] and Junaid Ali[2]

[1]Department of Electrical, Electronics and Communication Engineering, Engineering College, Ajmer, India
[2]Department of Mechanical Engineering, Indian Institute of Technology, Madras, India

ABSTRACT

Intelligent Speed Adaptive System (ISAS) is an emerging technology in the field of autonomous vehicles. However, the public acceptance rate of ISAS is drastically low because of several downfalls i.e. reliability and low accuracy. Various researchers have contributed methodologies to enhance the traffic prediction scores and algorithms to improve the overall adaptability of ISAS. The literature is scarce for Image Regression in this range of application. Computer vision has proved its iota in stream of object detection in self-driving technology in which most of the models are assisted through the complex web of neural nets and live imaging systems. In this article, some major issues related to the present technology of the ISAS and discussed new methodologies to get higher prediction accuracy to control the speed of vehicle through Image Regression technique to develop a computer vision model to predict the speed of vehicle with each frame of live images.

KEYWORDS

Intelligent Systems, Self-Driving Vehicle, Image Processing, Image Regression, Computer Vision, Automotive.

1. INTRODUCTION

With the rapid growth of automobile users across the world and the surplus adaption of state-of-art technology in the latest automobiles, the automotive industries have turned themselves from batch-type producers to mass vehicle producers. The growing user density in the region also increases the chances of jammed traffics, accidents, and nonetheless environmental pollution due to idling of hundreds of vehicles on the signal crossings. Researchers across the globe are working judiciously on each subject of traffic control systems, accident prevention systems, and pollution control systems. So many efforts have already been done in the direction of developing an Intelligent Speed Adaption System to control the speed of the automobile to prevent traffic and ultimately accidents due to high vehicle speeds, lack of control during cornering, and overtaking. With each passing generation of humans as well as automobiles, the dire need for high horsepower, more speed, cutting edge aerodynamic design, and intelligent behaviour of the vehicle in response of the driver is in demand and with Artificial Intelligence taking over the course of almost everything, automotive industries are also developing state-of-art intelligent vehicle systems to give premium experience to the customers [1].

ISA is based on a speed limiter incorporated within each vehicle that can take into account speed limit restrictions, that can adjust the maximum driving speed to the speed limit specified by the

David C. Wyld et al. (Eds): COMIT, CRBL, BIOM, WiMNeT, SIP, AISO - 2021
pp. 203-213, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.111718

roadside infrastructure, and that can provide feedback to the driver or take autonomous action when that speed limit is exceeded. ISA systems could use fixed or dynamic speed limits. In the fixed case, the driver is informed about the speed limit, which could be obtained from a static database. Dynamic speed limits take into account the current road conditions such as bad weather, slippery roads, or major incidents before prescribing the speed limit. If we assume Road Speed Limit as the range of speeds with a minimum and maximum value rather than a single absolute value. Then the highlighted benefit of such a system is that driver will have a low speed as well as high speed for an individual road to drive on. The selection of speed between the ranges will be defined as per mathematical algorithm fed on live traffic data in form of live image frames. In simpler words, the vehicle will choose to go slow if the traffic is dense on the road whereas the vehicle will choose to go fast but within the speed limit if the traffic is open on the road. However, the underlying problem with this method is limiting vehicle dynamics considerations in determining the road safety speed limits. The speeds assigned on the highways are determined based on vehicle dynamics of average automobile models which is not possible to be altered immediately. Such limitations are frustrating for drivers of the latest technology automobiles which can perform better because of advanced design and dynamics. We understand each vehicle has its defined performance capability a 1200 cc engine is going to over-take 800 cc engines for plausible engineering reasons. But it would be arrantly injustice for a 1200 cc engine vehicle to lag in traffic due to speed limits based on 800 cc engine capabilities. Revolutionizing the road speed limit data across the globe will be a difficult task. But we can revolutionize the way our vehicles respond to traffic and speed limits with the help of computer vision powered by image regression tools [2] [3].

Moreover, the vehicle speed prediction is not just limited to the above-mentioned factors but also the un-registered obstacles on the road e.g. wild animals crossing the road, construction work, pitfalls on road, and any unidentified object lying on road. To tackle this problem Computer Vision came to our rescue and with a dataset of 10,000 high-quality images provided by the Indian Transport Department. The dataset contains images of all possible obstacles a vehicle could face while driving on urban or rural roads. Computer Vision assisted with Convolution Neural Networks model envisages the vehicle to process live images from roads and make decisions backed by the mathematical algorithm to prevent a collision or slows down the speed of the vehicle on the road.

Most of the speed adaption system depends on GPS which are highly inaccurate and don't update from time to time, create lags while driving. A robust system is required that works offline and thus adapts the speed within seconds. Most of the GPS-based models are slow and thus cannot provide good results in urban areas. Here we have aimed to predict the speed of the vehicle using the images taken from the car dashboard. The neural net is been trained to take an input image segment it and thus provide a speed estimation. Each training image or frames speed is been annotated at a given time frame that is used to train the neural net. Thus, improving the overall accuracy of the computer vision model to detect obstacles and adjust speed based on algorithm [4].

2. Data Pre-Processing

At first, the image data is converted into a proper R, G, B channel with a height, width of 120x120x3. There are two types of data on which experiments is been done, Highway data and Urban data

For neural nets, the images are not been converted into arrays but for other models like SVR, Random Forest, and linear regression models the image is been converted into the NumPy arrays. At each frame, the speed of the vehicle is been recorded and thus annotated with frames. CNN-based image regression is mainly used to detect bounding boxes when there is a task of classification but if proper features is been extracted it can be used to predict the continuous values.

2.1. Background Subtraction in Image Processing

Background subtraction is a technique used to recognize the moving object in a video. The recognition of the images is done by using static cameras. The fundamental principle behind using background subtraction is recognizing the image from a difference between the reference frame of the image and the present frame of the image. The reference frame in this technique is known as the background image. The background model must be static while extracting the front objects from the video. The generic name of the front object in the video is known as the foreground object. Background subtraction is divided into two categories:

i. Parametric Background Subtraction
ii. Non-Parametric Background Subtraction

Two major techniques that are frequently used for background subtraction are pixel-based and block-based. In the case of statistical representation of the image non-parametric pixel technique is used. Shadow detection and illuminations are two major factors that impact the quality of the background subtraction. While capturing the foreground object, the background in the image must be static. The major application of background subtraction is in video surveillance and video analysis. The quality of the image extracted after the background subtraction can be enhanced by using various methods such as phosphine dots. These images as shown in figure 1 and figure 2 are then passed through neural nets and VGG 16.

Figure 1. Parametric Background Subtraction

Figure 2. Non-Parametric Background Subtraction

3. PROPOSED SOLUTION

When a proper CNN is constructed and while tuning the last layer to predict the continuous value rather than probabilities. While compiling the neural net MAE (Mean Absolute Error) with Adam optimizer is used. At here we proposed a solution to map a regression function using CNN for getting a prediction of speed from the images. For that, we have used the ISA^2 dataset [5]. Humans can't detect the speed of the vehicle from a single image but CNN can help to predict the same. That's why we proposed a novel solution to detect the speed of the vehicle using a single image. We proposed four algorithms for the same. The deep learning solution and three classical machine learning solutions.

3.1. Neural Network Regressor Model

The CNN regressor model directly takes an input image which is been rescaled in RGB format. The number of channels is not been reduced which passing the image from the CNN model. The input shape of the images been 120x120x3. The CNN regressor network is been trained to learn from the mapping from the input images with their labelled speeds.

$$\hat{s}_W = f(W, I_i) \tag{1}$$

The above equation hold for the mapping for the image to speed by CNN, where $f(W, I_i): I_i \rightarrow \mathbb{R}$ represents the mapping of the input image to speed. The loss function for the trainable model is been a mean square error (MSE).

Figure 3. Sample Dataset – Highway Traffic Images

Figure 4. Sample Dataset of Urban Traffic Images

These images as shown in figure 3 and figure 4 for the Highways and Urban data is been passed through CNN with the following properties.

The proposed neural net has an input layer with dimensions 120x120x3. The second layer consists of CONV2D with the size of 118x118x16 with a max-pooling layer of 59x59x16. Then second CONV2D layer is been inserted with the size of 28x28x32 with an additional global polling average and the dense layers. The output layer is accompanied by the MSE loss function. The model is been trained with an iteration of 1000 and validation error is been noted at each epoch. To evaluate the model MAE or mean absolute error has been used for both the image sets of Highway and Urban.

For K images in training or testing set, the MAE is given by equation 2 below,

$$\frac{1}{K}\sum_{i=1}^{K} \left|s_{r_i} - \hat{s}_i\right| \tag{2}$$

3.2. SVM Regressor Model

In SVM, the images can't directly feed onto the model, that why it needs to break down in array format, for this the images are been flatten in the dimension of 2800x43200 for training sets. The same dimension is been used for the urban image datasets. The images are been a break down in the arrays which contains the features of any given images. After that, the images are been normalize and passed through noise filters. The support vector regression model is trained using these images with labelled speed for each image set as shown in figure 5.

The SVR has the following constraints as shown in equations 3 to 7,

$$\min \frac{1}{2} \parallel w \parallel^2 + C\sum_{i=1}^{N} \xi_i + \xi_i^* \tag{3}$$

$$y_i - w^T x_i \leq \varepsilon + \xi_i' i = 1 \dots N \tag{4}$$

$$w^T x_i - y_i \leq \varepsilon + \xi_i i = 1 \dots N \tag{5}$$

$$\xi_i \xi_i \geq 0 i = 1 \dots N \tag{6}$$

$$\mathcal{L}(w, \xi', \xi, \lambda, \lambda^{\circ}, \alpha, \alpha^*) = \frac{1}{2} \parallel w \parallel^2 + C\sum_{i=1}^{N} \xi_i + \xi_i^* + \sum_{*1}^{N} \alpha_i^*(y_i - w^T x_i - \varepsilon - \xi_i^*) + \sum_{i=1}^{N} \alpha_i(-y_i + w^T x_i - \varepsilon - \xi_i) - \sum_{i=1}^{N} \lambda\xi_i + \lambda_i \xi_i' \tag{7}$$

Where the slack variable or the cost variable ξ_i is introduced.

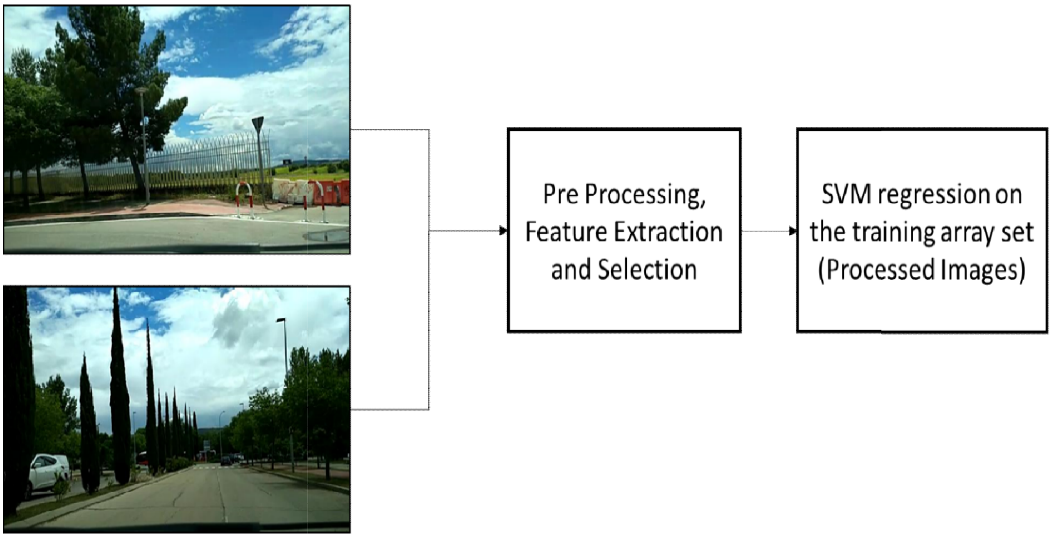

Figure 5. Algorithm of Feature Extraction and Speed Prediction

3.3. Random Forest Regression

The random forest regression model is a part of the supervised learning used to get higher prediction accuracy of traffic to control the speed of the vehicle. Numerous image training sets are assembled to improvise the higher prediction score. Random forest regression is a part of supervised learning based on the ensemble technique. The random forest technique can be used for both regression and the classification of the images. By taking the decision tree as a base of the model, column sampling and the rows sampling are done in the random forest technique. In the case of the random forest regression, if more image training sets are used then the variance of the model decreases, and hence the stability of the model increases. The random forest algorithm works by developing a decision tree multitude at the time of image datasets and provides the mean or mode value of all individual trees prediction values. Lower the value of the variance shows the higher stability of the model. When the image training sets are decreases then the value of the variance increases so, the stability of the model decreases. The mathematics behind the random forest regression is as shown from equation 8 to 12:

3.3.1. Splitting Criterion

$$RSS = \sum left\,(Y_i - Y_l)^2 + \sum right(Y_i - Y_r)^2 \tag{8}$$

Where,

$$Y_l = Left\,node\,Y - mean\,Value \tag{9}$$

$$Y_r = right\,node\,Y - mean\,value \tag{10}$$

3.3.2. Gini Criterion

$$Gini = n_L \sum_{n=1}^{k} p_{kl}\,(1 - p_{kl}) + n_R \sum_{n=1}^{k} p_{kr}\,(1 - p_{kr}) \tag{11}$$

Where,

$$p_{kl} = Left\ node\ proportional\ of\ class\ k$$
$$p_{kr} = right\ node\ proportion\ of\ class\ k$$

The Gini index value is used to determine the frequency at which any individual element among the dataset is mislabelled. The range of the Gini index is 0 to 1.

3.3.3. Mean Absolute Error (MAE)

The MAE in the case of the random forest regression model is given by

$$MAE = \frac{\sum_{I=1}^{N} abs\left(Y_I - \lambda(X_I)\right)}{N} \tag{12}$$

3.4. VGG 16 (Oxford Net)

VGG 16 is a CNN model used for image recognition at a very large scale. The dimension of the image training dataset used in this report is 2800 X 43200. The dimension of the input RGB image training dataset to convolution layer 1 is 120 X120 X3. In a VGG 16 architecture model, the input set of image datasets is allowed to pass through a series of convolutional layers. This stack of the convolutional layers comprises filters with tiny receptive fields. Receptive fields are used to capture the notions present in the images.VGG16 is part of the convolutional neural network used to solve the problem related to computer vision. Common computer vision problems are the classification of image datasets and regression problems. VGG model is divided into two categories based on the number of layers present in the model (VGG16, and VGG19). VGG is still a powerful classifier used for the classification of the image dataset. VGG16 algorithms can be used for regression and classification tasks. Mean squared value and the mean squared error in the case of VGG16 is given by the following mathematical relationship as shown in equation 13 and 14.

$$mae = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - X_i^{GT}\right| \tag{13}$$

$$mse = \left(\frac{1}{n}\sum_{i=1}^{n}\left|X_i - X_i^{GT}\right|\right)2 \tag{14}$$

4. RESULTS & DISCUSSION

To evaluate models and their performance the standard Mean Absolute Error metric is been used which computes the difference between the actual speed and predicted speed averaged over the K value of images. The results for the Highway as well as Urban areas is been calculated separately. The following tables is been constructed for the Highway as well as for Urban images.

Table I: Model Summary

Method	Highway (MAE) km/hr	Urban (MAE) km/hr
CNN based approach	10.45	9.640
SVM Regressor	8.473	6.460
Random Forest Regression	8.428	8.465
VGG16	9.668	8.279

From the above Table I, can be interpreted that the VGG16 has performed very well on the Highway data. Similarly, the SVM regressor has shown very good results in the urban image data. The tuned neural net has performed well on both datasets but didn't produce better results as compared to SVM or Boosted Trees.

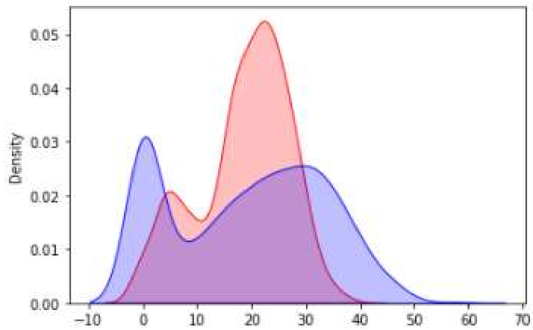

Figure 6. Density plot for predicted and actual values

The above plot shown in figure 6 describes the density plots for predicted and actual values where the red ones are of the predicted and the blue ones are for actual values. The predicted values show the same KDE pattern as the true images speed. This helps to determine how close the distribution of the predicted and the actual speed data. In figure 7 and 8, epochs are plotted against the losses and the decreasing trend is showing the increasing accuracy of the model which is desirable for accurate speed prediction.

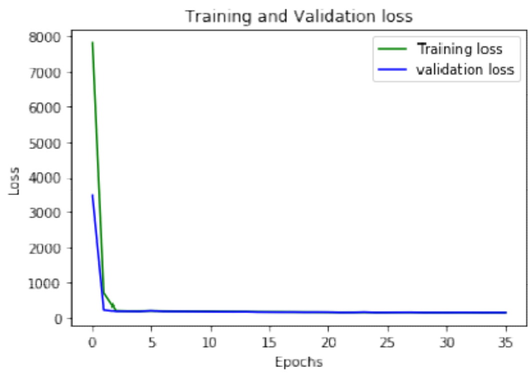

Figure 7. Loss -Epoch Curve - Highway Image Dataset

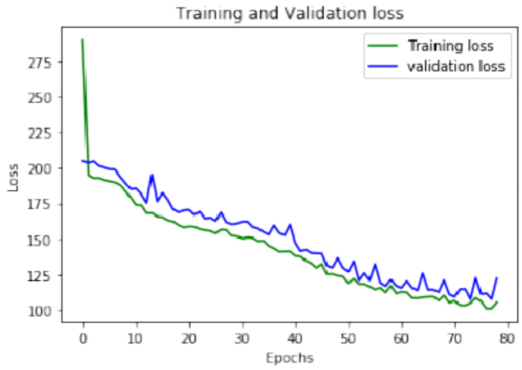

Figure 8. Loss- Epoch - Urban Image Dataset

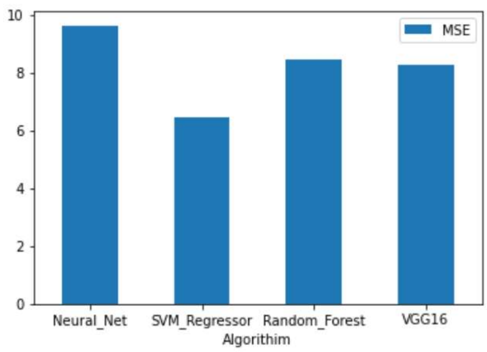

Figure 9. MSE comparison for Urban Data

The above results shown in figure 9and 10 the lowest MAE is been produced by the SVM regressor on the Urban data. The SVM regressor has performed better than the VGG16 and Neural Net as the MSE of these two is very high. The Urban data don't have many features as the roads and traffic density is very less as compared to Highway Data.

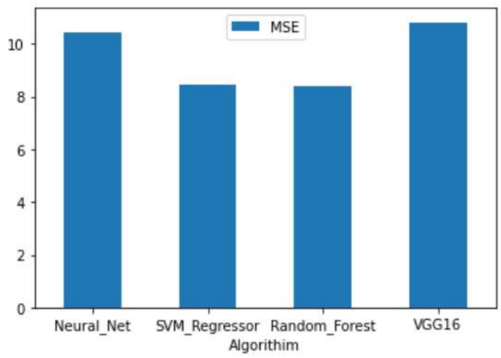

Figure 10. Comparison of MSE in Highway data

The number of features in the High way data is much more as compared to the Urban Data as the number of cars and traffic density increases in Highway data. SVM regression and Random forest have worked very well on the Highway data as compared to VGG16 and Designed Neural Net.

5. CONCLUSION & FUTURE SCOPE

In this article we have performed computing experiments to conclude that through the given image frame from vehicles while driving, the speed of the vehicle can be detected using machine learning algorithms. Our objective was to demonstrate the capability and prowess of Image Regression in computer vision techniques to detect the speed of running vehicle which is potent as compared to the conventional GPS tracking APIs and vehicle platooning as previous works claims. We have achieved our objective by comparing different machine learning and neural networks models on same dataset. Among the Highway Image data and Urban Image data theleast MAE has been shown by the SVM and Random Forest Regressor whereas VGG 16 and NN have shown comparatively poor results. The ML algorithms were able to capture the features and thus able to make a speed predictions with minimal error. In future a segmentation technique and optical flow can be used enhance the model performances through reducing the computing time through hyper parameters tuning and advance algorithms. Image regression for speed detection of vehicle can be interlocked with automobile ECU to control braking and speed control in dense traffic areas and accident prone areas. Computer vision assisted machines have better visibility range and reaction timing than a human beings, wide angle and multipoint focus is far superior then human retina. In countries like India, Pakistan and Bangladesh where roads are flooded with wildlife and carefree human beings, intelligent decision making can save many human and animal lives by preventing road accidents.

6. ACKNOWLEDGEMENT

We would like to show our gratitude towards IROS 2018 Workshop, 10th Planning, Perception and Navigation for Intelligent Vehicles for providing us with valuable dataset for our research.

REFERENCES

[1]. L Baskar, B De Schutter, and H. Hellendoorn, “Intelligent speed adaptation in intelligent vehicle highway systems – A model predictive control approach,” Proceedings of the 10th TRAIL Congress 2008 – TRAIL in Perspective – CD-ROM, Rotterdam, The Netherlands, 13 pp., Oct. 2008.

[2]. D Shukla and E Patel, “Speed determination of moving vehicles using lucas-kanade algorithm,” IJCATR, vol. 2, pp. 32–36, 01 2012.

[3]. I Sreedevi, M. Gupta, and P. Asok Bhattacharyya, “Vehicle tracking and speed estimation using optical flow method,” International Journal of Engineering Science and Technology, vol. 3, 01 2011.

[4]. A.J.E Atkociunas, R. Blake and M. Kazimianec, “Image processing in road traffic analysis,” in Image Processing in Road Traffic Analysis, vol. 10, 2005, pp. 315–332.

[5]. Herranz-Perdiguero, C. and López-Sastre, “ISA2 : Intelligent Speed Adaptation from Appearance”, IROS 2018 Workshop, 10th Planning, Perception and Navigation for Intelligent Vehicles (PPNIV' 18)