Tracing of Vehicle Region and Number Plate Detection using Deep Learning

Shantha Lakshmi S
computer science and engineering
sona college of technology
Salem, India
sweetshanthu95@gmail.com

Sathiyabhama B
computer science and engineering
sona college of technology
Salem, India
sathiyabhama@sonatech.ac.in

Revathi T K
computer science and engineering
sona college of technology
Salem, India
revathi.tk@sonatech.ac.in

Basker N
computer science and engineering
sona college of technology
Salem, India
baskern@sonatech.ac.in

Vinothkumar R B
computer science and engineering
sona college of technology
Salem, India
vinothkumarrb@sonatech.ac.in

2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) 978-1-7281-4142-8/20/$31.00 ©2020 IEEE 10.1109/ic-ETITE47903.2020.357

***Abstract*—Excessive booming global population creates a widespread increase in traffic on Hi-tech modern roads. It is very much essential to detect vehicles and computing traffic congestion on highways. The key aspect is to hoard vehicle data through number plate detection. A deep learning model or Network has been developed to detect vehicles passes dynamically and efficiently. In the proposed work, a deep learning-based algorithm is proposed for detecting both vehicles and number plates for a reputed company surveillance dataset. The proposed model uses a video dataset as an input and the video has been segmented into several frames. Using pre-trained weights and labels of the dataset, the vehicles and its number plate are detected by the dark flow toolkit. This tool provides to extract the region of the vehicle with proper annotation. In future work, the proposed model aims to calculate the speed of the vehicle based on the surrounding area.**

Keywords—Dark flow, Deep learning, Number plate detection, Vehicle Detection.

I. INTRODUCTION

Excessive booming of the global population creates a widespread increase in traffic on Hi-tech modern roads. It is much needed to detect the vehicles and computing traffic congestion. The application of number plate detection could be directing the identification of vehicles, giving alerts and warning, speed control according to surrounding areas.

One of the popular artificial intelligence techniques [3] to work with large amount of data is machine learning. It is a self-learning algorithm. It accomplishes with improvised analysis and patterns [3]. Whereas, deep learning, a sub class of machine learning uses a hierarchical level of neural networks [4-5] to carry out the process of machine learning. The artificial neural networks [2] perform like a human brain and neuron nodes connected perform as the link.

A. Neural Network

Deep learning is often known as deep neural networking or learning [4]. Neural networks may emerge in different forms like convolutional, artificial, recurrent and feed forward neural networkings. Each is based on different use cases but all functions in more or less in some ways, by collecting input field and enhancing the network to calculate for itself whether to analyze the decision made is right or wrong about the given input. The Neural network is a trial-and-error process [4], so it needs a large amount of data for training the model with better results.

B. Activation Function

Activation function [2] identifies the activation of neuron. It is calculated through the total sum and bias added to it. The idea of activation function is to include the output of a neuron with non-linearity. The variations of activation function include: (i) linear function is operated only in one place. (Output layer). (ii) Sigmoid function is operated with binary classification present in output layer, where the result can be of either 0 or 1, as value lies between 0 and 1. (iii) Tanh function is operated in hidden layers of a network which value is between -1 to 1 and it is mainly used for centering the data. (iv) ReLu function [5] involves simpler mathematical operations and only a few neurons are activated at a time to make the network improvised and easy for prediction. Softmax function is used to handle multiple classes and optimizes the output for each class between 0 and 1.

C. Deep Learning

Deep Learning being the sub class of machine learning algorithms [1] that has prompted by structural and functional part known as artificial neural network. The algorithm enhances the multiple hidden layers to spilt higher level features of the input. Most of the deep learning models apply artificial neural networks. The word "deep" in "deep learning" [1] refers to the number of layers applied for data transformation. The model has achieved remarkable results in computer vision, speech recognition and image classification [7]. Many models are implemented for image classification systems like detecting faces, identifying people in images, recognizing facial expressions [9] and gesture identification. It helps to abstract and select features to improve the efficiency of the system.

II. RELATED WORKS

The Google Inception Net (GoogLeNet) architecture, Shengyu Lu. et al. (2018) use the YOLO network. The network utilizes an optimal convolution function to modify the original convolution function, in order to minimize the number of arguments and for minimal detecting time [1].

978-1-7281-4142-8/$31.00 ©2020 IEEE

Besides the other related algorithms, the Fast YOLO and YOLO has an improvised precision values.

Detection of car license plate problem using a convolutional layer was addressed by Hendry. et al. (2019). The detector uses the sliding window to solve small objet detection issues in each image for all classes. The advantage of the sliding window -YOLO was their simultaneous phase of both detection and recognition. The carried from auto localization to character grouping of the plate area.

A novel multi-task framework for object detection, named as MONet [3]proposed by Tao Gong. et al. (2019) was the first framework which gives a simultaneous solution for both the multi-label classification and the object detection task.

Based on the technique to detect, localize and recognize actions of interest, Shubham Shindea. et al. (2018) used a real-time video data which is collected from a surveillance camera [4]. They also concluded that a single frame is enough for action recognition for some cases of their dataset.

Automated vehicle detection and counting system in aerial images proposed by Hilal tayara. et al. (2018) utilizes a convolution neural network [5] to regress a vehicle spatial density map across the aerial image. They concluded with the highest true positive rate and lowest false alarm rate.

A fast deep neural network [6] for real-time video object detection by wenming cao. et al. (2018) explores the ideas of knowledge-guided training and predicted regions of interest. It improves the performance of network and also minimizes the overall complexity for computational functions.

A video-based vehicle counting framework [7] proposed by Zhaoyang zhang. et al. (2018) used a three-component process of object detection, object tracking, and trajectory processing to obtain the traffic flow information. They also discussed the framework which is capable of vehicles count in all categories and route direction with an accuracy of 90% and more.

Integration of additional prediction layers into conventional Yolo-v3 developed by Kwang-ju Kim. et al. (2019) used spatial pyramid pooling [8] to complement the detection accuracy of the vehicle for large scale changes. They have also discussed that the proposed architecture shows a mAP detection ratio against the other vehicle detection with improvised speed of run-time.

III. Background

The implementation of vehicle and number plate detection is done through the YOLO and the toolkit used is Dark Flow.

A. You Only Look Once

You Only Look Once (YOLO) is a real-time object detection system. A single neural network is applied to the whole image [10]. This method splits the single frame into regions and then the boundary box is predicted. The predictive values detect the object by classification models. It predicts with a single network evaluation extremely faster than existing methods. Multiple objects can also be detected in a single frame. YOLO algorithm is faster; it is used to detect the vehicle region.

B. Dark Flow

There are different tools to implement YOLO. One such toolkit used to implement this model is Dark Flow. It takes video as an input and an extended version of darknet with Tensorflow. It includes collecting the dataset, annotation, configuring files, training and validating the dataset.

IV. Proposed System

A. Detecting the vehicle in the road

The vehicle detection is a very important component in traffic surveillance and to solve the increasingly serious traffic problem [8]. The detection of number plates is a tedious process. To overcome this problem, a two-step process is developed. The detection of vehicle region is a first step and the number plate of each vehicle is traced as a next step. This two-step approach reduces the search of number plate detection in a frame and it makes easy to detect number plate regions.

Vehicle Classification

Vehicle classification is a process of hoarding information about any of a detected vehicle's categories like the image containing a car, motorbike, bus or image without any vehicle. Now the vehicle's position [2] in the image is classified. The detection is done by bounding the object with a rectangular box according to the surrounding area.

Boundary Box

Besides vehicle classification, the proposed model needs to identify the position of the vehicle in the frame. This is possible by fixing a bounding box. A bounding box is generally represented by the center (b1, b2), rectangle height (b3), and rectangle width (b4). The model is trained with these four variables (b1, b2, b3, and b4) and assigning those values to the center box (b1, b2) [3]. After the neural network is trained with this labeled data, the model predicts the four variables values with colored bounding boxes.

B. Detecting the number plate in the detected vehicles

Vehicle number plate detection is processed by feature identity detecting process. The detected vehicle areas identified during the vehicle detection step are utilized in this stage. The identified vehicle detects an area having characteristics similar to vehicle region such as to determine a position belonging to the next frame. The successive frame fixes the detection area by the previous frames with detected vehicle region. Then, the number plate in the vehicle region can be recognized [5] and detected in the consecutive frames.

C. Predicting the multi object in the frame and alert for overlap with vehicle

Single Object Detection

YOLO algorithm performs a single neural network [10] for the entire input image. The model splits the region into the S*S grid. For each object, boundary boxes are marked and probabilities for each region are predicted. For probability prediction, logistic regression is applied. Logistic regression is the predictive analysis for the dependent variable and its relationship. Each object is predefined with a labeled and threshold value in the training dataset [11]. For example, the image is divided into 3*3 grids and contains three classes in which the object may be classified as a car, motorbike, and bus respectively. For each grid cell, the labeled value will form a vector.

$$X = \{P, B1, B2, B3, B4, C1, C2, C3\} \quad (1)$$

The value of P in (1) represents the probability of an object, B1, B2, B3, B4 represents the values of boundary boxes and C1, C2, C3 represents the classified classes.

Multiple Object Detection

The system also includes anchor boxes. Anchor boxes are with a defined aspect ratio; they detect objects that nicely fit into a box [13] with that ratio. Consider that the objects are overlapped and the anchor boxes are used. For example, overlapped objects, car and bus is predicted. One anchor box will define the shape of the bus whose width is greater than its height and another anchor box will define the shape of the car whose height is greater than its width. And the two predicted values are obtained. The overlapping objects can be predicted by increased S*S grid values (say 19*19 and more). After the output vectors are collected, this algorithm uses Non Maximal Suppression (NMS) for eliminating boundary boxes. The first part of NMS is to remove the predicted boundary boxes that have a value lower than the given thresholds (predefined). The next part in NMS is to highlight the bounding boxes with the highest detection probability and remove all the bounding boxes that have a value higher than the given Intersection Over Union (IOU) threshold. It is repeated until all of the non maximal boundary boxes are removed.

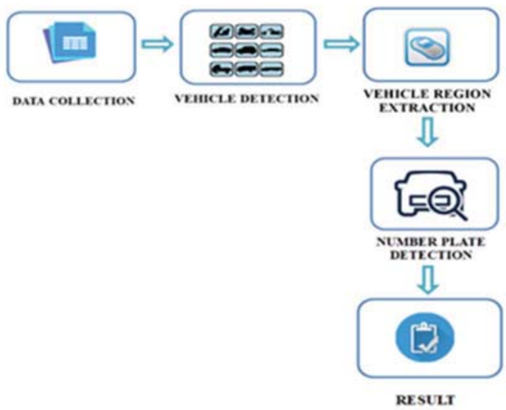

Fig. 1. Design of Number Plate Detection

V. IMPLEMENTATION

A. Datasets

The dataset was obtained from the T.A.Perumal Starch industries surveillance dataset. The dataset is of video format which is converted into frames for processing. The training and test dataset consists of vehicle and non vehicle data which is labeled as a vehicle or non vehicle in the COCO dataset. The dataset comprises of 9000 images of size (64x64 pixels). The author splits the model with 7000 images for training and 2000 images to test the dataset. The vehicle dataset contains the vehicle's front view, back view, and side view. The dataset also contains the non vehicle data that includes person, traffic light, etc. The input data is fed into the feature extraction method which generates the output data that has to be fed into the proposed model.

VI. PERFORMANCE EVALUATION

The proposed system evaluates the performance metrics of the vehicle number plate detection algorithm by comparing the ground truth data and then check if there exists values of confused matrices.

- TP = True Positive (overall vehicles identified by the model)
- FP = False Positive (overall vehicles identified by the mode even when it does not belong to)
- FN = False Negative (overall vehicles ignored by the model but it belongs)
- TN = True Negative (overall vehicles correctly rejected by the model)

Metrics used for performance evaluation are the standard methods such as Precision, Recall, and Accuracy.

$$Precision = TP/(TP + FP) \quad (2)$$

$$Recall = TP/(TP + FN) \quad (3)$$

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \quad (4)$$

The dataset obtained from T.A.Perumal Starch Industries surveillance is trained with the proposed model. The performance evaluation of our dataset is compared with other network models and obtained with better results.

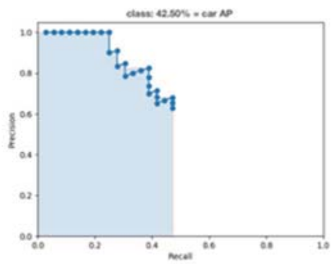

Fig. 2. Average Precision of proposed model

VII. RESULT

The YOLO was implemented on dataset collected from the company surveillance dataset for vehicle detection in the roads. The dataset split is in the ratio of 4:1. The training set with 80% images and 20% images for testing dataset and the performance metrics are tabulated. And the results are compared with other models that are tabulated for each class.

TABLE I. PERFORMANCE METRICS FOR DATASET TRAINED ON YOLO AND CNN MODEL

Average Precision	Frames per Second (fps)	Algorithm	
		YOLO	*CNN*
Class : Car	45	42.50%	41.50%
Class : MotorBike		42.86%	42.1%

REFERENCES

[1] Shengyu Lu , Beizhan Wang , Hongji Wang , Lihao Chen , Ma Linjian , Xiaoyan Zhang "A real-time object detection algorithm for video", Computers and Electrical Engineering 77 (2019) ELSEVIER 398–408.

[2] Hendry, Rung-Ching Chen "Automatic License Plate Recognition via sliding-window darknet-YOLO deep learning", Image and Vision Computing 87 (2019) ELSEVIER 47–56.

[3] Tao Gong, Bin Liu , Qi Chu, Nenghai Yu "Using multi-label classification to improve object detection", neurocomputing, ELSEVIER(2019).

[4] Shubham Shinde, Ashwin Kothari, Vikram Gupta "YOLO based Human Action Recognition and Localization", ELSEVIER (2018).

[5] Hilal tayara, Kim gil soo, and Kil to chong "Vehicle Detection and Counting in High-Resolution Aerial Images Using Convolutional Regression Neural Network", Digital Object Identifier 10.1109/ACCESS.2017.2782260.

[6] Wenming cao , (member, ieee), Jianhe yuan, zhihai he , (fellow, ieee), Zhi zhang, and Zhiquan he (member, ieee) "Fast Deep Neural Networks With Knowledge Guided Training and Predicted Regions of Interests for Real-Time Video Object Detection", Digital Object Identifier 10.1109/ACCESS.2018.2795798.

[7] Zhe Dai , Huansheng Song, Xuan Wang, Yong Fang , Xu Yun, Zhaoyang Zhang, And Huaiyu Li "Video-Based Vehicle Counting Framework", Digital Object Identifier 10.1109/ACCESS.2019.2914254.

[8] Kwang-Ju Kim , Pyong-Kun Kim, Yun-Su Chung, And Doo-Hyun Choi "Multi-Scale Detector for Accurate Vehicle Detection in traffic surveillance data"digital object identifier 10.1109/access.2019.2922479.

[9] Ajeet Ram Pathak, Manjusha Pandey, Siddharth Rautaray "Application Of Deep Learning for Object Detection", elsevier (2018).

[10] YOLO real-time object detection https://pjreddie.com/darknet/yolo/.

[11] https://towardsdatascience.com/an-introduction-to-implementing-the-yolo-algorithm-for-multi-object-detection-in-images-99cf240539.
Introduction to YOLO implementation.

[12] Dhiraj Sangwan, Deepak Kumar Jain, "An Evaluation of Deep Learning Based Object Detection Strategies", pattern recognition letters (2019),doi: https://doi.org/10.1016/j.patrec.2019.01.014.

[13] Geethapriya. S, N. Duraimurugan, S.P. Chokkalingam "Real-Time Object Detection with Yolo", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-3S, February 2019.