Brief Paper:
Vehicle Manufacturer Recognition using Deep Learning and Perspective Transformation

Israfil Ansari[1], Jaechang Shim[1*]

Abstract: In real world object detection is an active research topic for understanding different objects from images. There are different models presented in past and had significant results. In this paper we are presenting vehicle logo detection using previous object detection models such as You only look once (YOLO) and Faster Region-based CNN (F-RCNN). Both the front and rear view of the vehicles were used for training and testing the proposed method. Along with deep learning an image pre-processing algorithm called perspective transformation is proposed for all the test images. Using perspective transformation, the top view images were transformed into front view images. This algorithm has higher detection rate as compared to raw images. Furthermore, YOLO model has better result as compare to F-RCNN model.

Key Words: Vehicle Logo, Object detection, YOLO, Faster R-CNN, VMR.

I. INTRODUCTION

In recent years many object detection and classification algorithms are proposed. By using those algorithms, anyone can train their custom object for realtime detection. Human can instantly identify the object in an image at one look, but deep learning algorithm fails in many situations.

Vehicle logo is one of the most essential highlighted descriptors, which can furnish the astute activity framework with valuable data to distinctive vehicles. In past the Vehicle Manufacturer Recognition (VMR) was done using License Plate Recognition (LPR). But due to low picture quality LPR systems fails in different situation and have many detection errors due to the problem of character segmentation of license plate. Vehicle manufacturer recognition is getting more and more complex due to increase of different vehicle models in the market. The main difficulty of VMR is that there are many vehicle models today of almost same design and changing quickly overtime. Therefore, detection of the vehicle logos from various manufacturers is chosen as the main feature of the vehicle. VMR is a field with limited research, among which vehicle logo recognition [1] is widely used because logos are the unique point for the entire manufacturer.

In past many researchers applied different image featured algorithms and models but many of them were not robust in some complex images. In [2] a SIFT based method was proposed for vehicle logo detection and recognition. The system was enhanced by merging features using multiple images. They also used generalized Hough transformation for feature clustering and for affine transformation a generic verification was applied. In [3], the authors used 1200 logo images of 10 distinctive vehicle manufacturers to assess a SHIFT- based approach. Than an enhanced feature matching method was used by merging SIFT points from the provided sample logo. The accuracy presented in this method was 97%, however, the test images were close-up logo images and does not corresponds to street traffic cameras.

II. PROPOSED SYSTEM

In this study we are proposing advance VMR using deep learning. Here, two deep learning algorithms Faster R-CNN [4] and YOLO [5] is compared. This system is divided into two parts: (i) training data (ii) testing data. In training part, we have collected sample images of different vehicle manufacturer logos. In both deep learning algorithms same images were used while training. The training architecture is depicted in Fig.1.
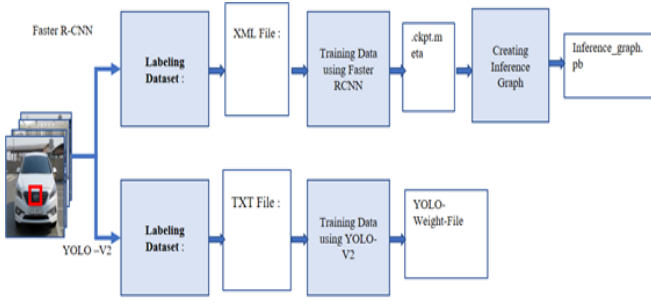
Fig. 1. Training architecture of proposes system.

The training architecture of both Faster R-CNN and YOLO are shown in Fig. 1. All the sample images were labelled according to the class of vehicle manufacturer. We have labelled four classes, Hyundai, Kia, Samsung, SsangYong, Daewoo. After labelling, we have trained the image data with both the deep learning models.

Both the trained models were saved for testing images captured from traffic CCTV cameras. Before testing any images, we propose to pre-process each image. For image pre-processing we used advance Perspective Transformation (PT) [6] method through which we make the logo more distinct. PT is proposed because when the images are captured from traffic cameras the logos are not visible for all vehicle model. There is some brand where logos are installed on the bottom of the vehicle's body. For perspective transformation, a 3x3 transformation matrix is required for processing. For matrix, 4 points are required from the input image and corresponding points. Direct linear transformation algorithm is being used to transform) to   with projection matrix M.

$$X_i^/ = MX_i, X_i^/ = MX_i,$$ (1)

$$t \begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$ (2)

where t is the weight of the feature points. Equation (1) is the scalable variable. We can derive equation (3) from equation (2):

$$x_i' = \frac{m_{11}x_i + m_{12}y_i + m_{13}}{m_{31}x_i + m_{32}y_i + 1} \quad , \quad y_i' = \frac{m_{21}x_i + m_{22}y_i + m_{23}}{m_{31}x_i + m_{32}y_i + 1}$$ (3)

In projection matrix M, we must find 8 elements. If we know 4 pairs of source and destination point, then we can find projection matrix M.

After PT, the transformed images are then used to find the logo of the manufacturer. Fig. 3 shows the architecture of testing process; all the input images are transformed initially before sending it to deep learning algorithm to detect and recognize vehicle manufacturer.
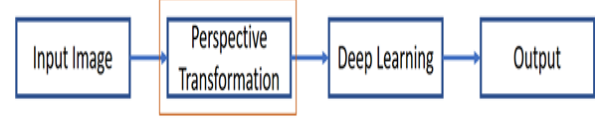

Fig. 2. Test Architecture of proposes system. Input image is the vehicle images, after getting the input images PT is applied on those images. After PT it is then forwarded to the pre-trained model of deep learning for VMR.

III. RESULT AND DISCUSSION

Traffic CCTV camera are generally installed on the top of the road therefore an area of interest must be defined to capture vehicle images . If the car images are taken from location Fig. 4(a), the picture quality is low therefore it's very difficult for VMR. Vehicle images from Fig. 4(c) is viewed from top, so it is difficult to track the logo. We propose the location shown in Fig 4(b). From this location both front view logo and rear view logo of vehicles have higher detection rate after PT.
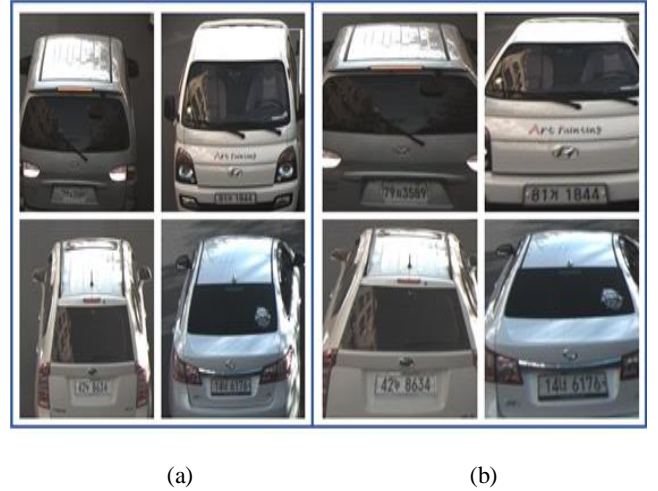

(a)   (b)

Fig. 3. Perspective Transformation, (a)Images taken from traffic CCTV camera, (b) Images after Perspective Transformation.

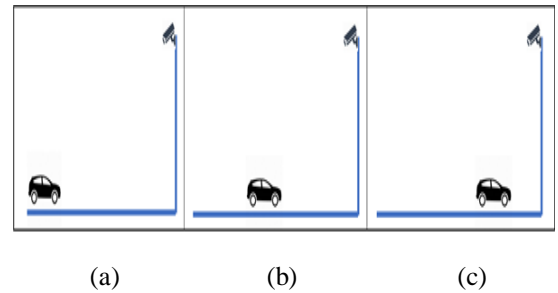(a)   (b)   (c)

Fig. 4. Vehicle image capture area. (a) Area where vehicle appear on the camera frame, (b) Middle of the camera frame, (c) Top view area from the camera.

The proposed method is implemented on system with an Intel Core i7 CPU, 16GB memory, and GTX 2080 GPU.

Our training dataset consist of classes i) Hyundai with 500 images, ii) Kia with 300 images, iii) Samsung with 300 images, iv) SsangYong with 250 images v) Daewoo with 100 images. All the training data set were combination of front facing camera and few from area shown in Fig. 4(a). All the images were labelled and trained with both the deep learning algorithms.

After training, the trained models are used to test data received from traffic CCTV from the view location Fig. 4(b). To test our algorithm, we have collected vehicle images during day light and night of different vehicle brands as shown in Table 1. As shown in Fig. 5 and Fig. 6, Method 1 is the result of Faster R-CNN without perspective transformation of vehicle images. Method 2 is the result of Faster R-CNN after perspective transformation of vehicle images. Method 3 is the result of YOLO algorithm without perspective transformation of vehicle images. Method 4 is the result of YOLO algorithm after perspective transformation of vehicle images.

Table 1. Total test images collected from street CCTV.

Vehicle Manufacturer	Total Day Images	Total Night Images
Hyundai	200	186
Kia	50	19
Samsung	50	14
SsangYong	50	10
Daewoo	10	3

VMR is essential in many sectors, and many algorithms has shown significant results. When VMR is used for the images taken from a traffic camera, it is difficult to recognize due to various reason. As, shown in Fig. 5, method 1 has a good result but the result after PT has enhanced significantly in method 2. Furthermore, method 3 has a better result than method 1 and 2, but when PT is applied in method 3, the result has a drastic change.

As shown in Fig. 5, detection rate of Hyundai, Samsung, SsangYong in method 4 has reached to 100 percentage, whereas Daewoo has increased from 20 percentage in method 1 to 80 percentage in method 4. In night test data as shown in Fig. 6, the result has significantly changed from method 1 to method 4. In Table 2 all the results in percentage are shown from method 1 to method 4 on both day and night test data.
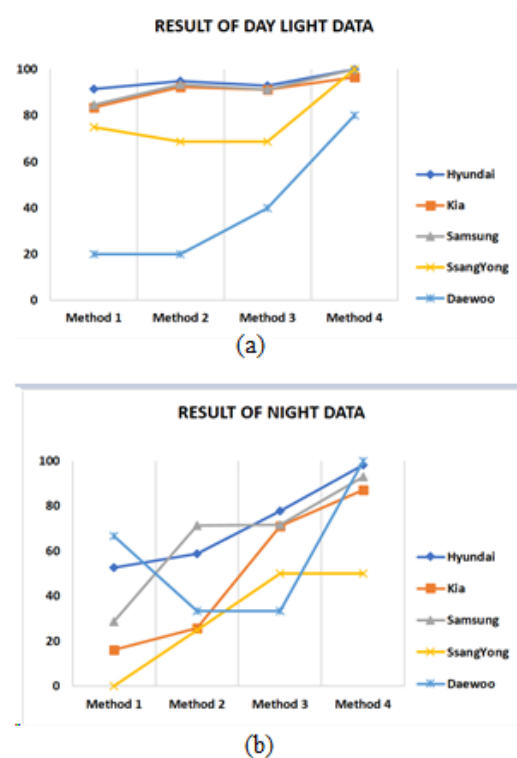

Fig. 5. Result of day and night data, (a) Result of proposed algorithm on day light vehicle images. As seen in the figure, using PT the detection rate has increased as compared to without PT. Furthermore, YOLO-V2 after PT has better result than Faster R-CNN after PT, (b) Result of proposed algorithm on night vehicle images. As seen in this figure, here also the detection rate has increased using PT. YOLO-V2 has better result as compared to Faster R-CNN. Detection rate of the classes has increased from very low to nearby around 90 percentages.

Table 2. Result of VMR.

Vehicle Model	Method 1(%)		Method 2(%)		Method 3(%)		Method 4(%)	
Vehicle Model	Day	Night	Day	Night	Day	Night	Day	Night
Hyundai	91.4	52.8	95	58.7	93.1	77.6	100	98.0
Kia	83.5	16.1	92.3	25.8	91.2	80	96.7	87.1
Samsung	84.8	28.6	93.5	71.4	91.3	71.4	100	92.9
SsangYong	75.0	0.0	68.8	25.0	68.8	50.0	100	50.0
Daewoo	20.0	66.7	20.0	33.3	40.0	33.3	80	100

IV. CONCLUSION

This proposed system has compared two well-known object detection and recognition algorithm for vehicle manufacturer recognition. These object detection algorithms have good results but failed in some cases while detecting small objects such as vehicle logo from images captured from traffic cameras. To overcome the problem, we propose to apply perspective transformation algorithm to the images before applying it to deep learning algorithm. We also propose an area from where vehicle images need to capture. Among all the four methods, our proposed method 4 has best results on both day and night images.

There are some issues with night data as the picture quality is very low. In future, we plan to increase the picture quality of the images to detect and recognize vehicle logos.

REFERENCES

[1] S. Mao, M. Ye, X. Li, F. Pang, J. Sun, "Rapid vehicle logo region detection based on information theory," *Computers and Electrical Engineering*, pp. 863-872, 2013.

[2] A. Psyllos, C. N. Anagnostopoulos, and E. Kayafas, "M-SHIFT: A new method for Vehicle Logo Recognition," in *Proceeding of IEEE Inter. Conf. on Vehicular Electronics and Safety*, pp. 24-27, 2012.

[3] A. Psyllos, C. N. Anagnostopoulos, and E. Kayafas, "Vehicle Logo recognition using a shift-based enhanced matching," *IEEE Trans. on Intell. Transp. Sys.* vol. 11, pp. 322-328, 2010.

[4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017.

[5] R. Joseph, F. Ali, "YOLO9000: Better, Faster, Stronger," https://pjreddie.com/publications/yolo9000/, accessed August 2018

[6] T. Shakunaga, H. Kaneko, "Perspective angle transform: Principle of shape and angles," *International Journal of Computer Vision*, vol. 3, pp. 239-254, 1989.