Event Analysis for Vehicle Classification using Fast RCNN

Khaing Suu Htet
GIS Research Lab
University of Computer Studies,Yangon
Yangon, Myanmar
khaingsuuhtet@ucsy.edu.mm

Myint Myint Sein
GIS Research Lab
University of Computer Studies,Yangon
Yangon, Myanmar
myintucsy@gmail.com

***Abstract*— Event analysis on object detection is critical as classification of vehicles and persons in events. This system proposed vehicles classification and counting for event system. The images from event video streaming camera are first featured extracted, segmentation, model building and presented to achieve improving optimization by Deep Neural Learning Fast R-CNN method with hyper-parameter optimization using the two datasets, modified Stanford car dataset and new Myanmar cars dataset. Experimental result demonstrated that Fast RCNN and hyper-parameter classifier with created Myanmar cars dataset attained the state-of-art precision result in vehicle classification and counting on real-life event Video streaming.**

Keywords*— *vehicle classification and counting, Fast-RCNN with parameter optimization, Myanmar Cars Inage Dataset, event traffic analysis, hyper-parameter optimization

I. INTRODUCTION

Traffic management and car parking arrange is challenging factors from a history of working in event or festival. Classification and Counting Vehicles is critical area for Event Analysis. Vehicles and mobile devices moving in and around, reversing, loading and unloading at an event site can create a serious risk of incidences. Vehicle classification and counting is used as technique for better analysis and management of events as religious ceremony, recreational, social, sporting and fundraising. Image processing is a critical role in detecting vehicles from Event surveillance monitoring video streaming. Event traffic can classify as buses, cars, motor bikes, trucks. Vehicle detection has achieved video streaming from many surveillance cameras set up entrances of event.

The problem is well-regarded as object detection and classify the images containing vehicles into the right type and it has challenge in the wrongly recognized small types of vehicles.[1] Song, Huansheng discussed the multi-object tracking and trajectory analysis with large-scale high datasets rather than the public datasets and it applied CNN and ORB algorithm to sum up the car movement direction and the traffic flow on highway traffic scene. [2]

CNN has applied in many vehicle detection in Europe. Fast RCNN was applied for vehicle detect and classification in traffic areas in the city of Karlsruhe, Germany.[3] Fast RCNN applies a selective search strategy to differentiate all candidate frames, which is significantly time-consuming, and the speed of detection vehicle is slow. The increasing complexity of fast RCNN methods arises many issues, the dominant one is "overfitting" and this occurs in the training dataset process, but is not able to detect significant different objects outside of the dataset.[4] This study targets on developing high accurate classification task with improving the recognition rate, i.e. percentage of correctly classified images. The result is influenced not only using separate datasets, but also improving additional training parameters that called hyper-parameters. To get optimal hyper-parameters values, Nelder-Mead method is applied to hyper-parameter analysis learning environments. This optimization search method will help to achieve cost-effective time saving comparing with others AI-powered search intelligence.

In this system, vehicle classification and count on video streaming from Myanmar wedding event is applied with hyper-parameter optimization on fast R-CNN. The proposed system also prepared and trained using hyper-parameter on Stanford Car Dataset and Myanmar Car Dataset. This paper includes 5 Sections, Section-II describes the system overall landscapes, and proposed method as fast RCNN and hyper-parameters with Nelder-Mead optimization. And Section-III includes how to train two datasets by using Fast RCNN and additional hyper-parameter. Section-IV describes the experimental result and finding based on real-life wedding event streaming video and succeeding by conclusion in next Section-V.

II. SYSTEM LANDSCAPE

This system has used Fast RCNN model and it is updated with hyper-parameter optimization to classify Vehicles and counting as Fig.1.

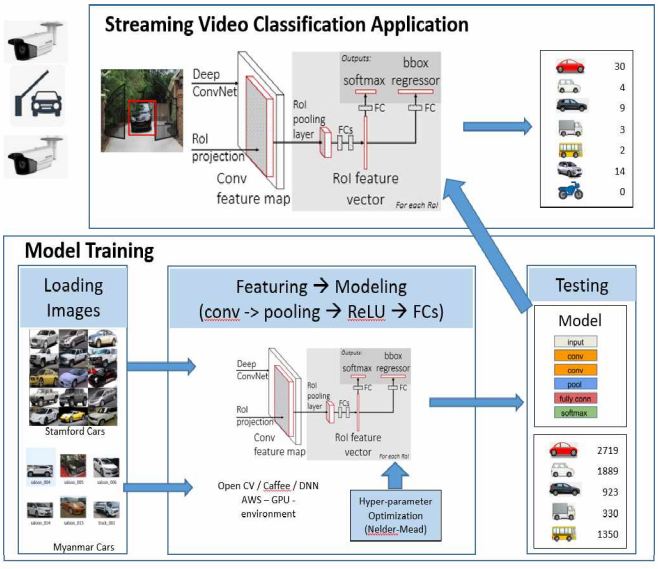

Fig.1. Overall System Landscapes

978-1-7281-9802-6/20/$31.00 ©2020 IEEE

A. Object Extraction and Classification (Fast RCNN)

Images from train data and video streaming are detected and extracted. R.Girshick proposed Fast RCNN to reduce computing time in R-CNN algorithm and it consist a backbone CNN and final pooling layer is take out an "ROI pooling" layer and FC layer is stand in (a(K+1) softmax layer and bounding box regression. The algorithm extract feature from every regions and it divides into different classes, and achieve the boundary boxes for the specific class continuously. [5] In all training network, weight values with back-propagation is important for Fast R-CNN. This paper used efficient training with adding hyper-parameter as Table I.

TABLE I. SOME HYPER-PARAMETERS OF TRAINING MODEL

Name	Description	Range
$x1$	Learning rate (= 0.1$x1$)	[0, 4]
$x2$	Momentum (= 1 - 0.1$x2$)	[0.5, 2]
$x3$	L2 weight decay	[0.001, 0.01]
$x8$	Conv 1 initialization deviation	[0.01, 0.05]
x14	Conv 1 bias	[0, 1]

B. Hyper-parameter Optimization Model

The Nelder-Mead method is simple and applied hyper-parameter tuning problem by Nelder and Mead. Gilles et al. with support vector machine modeling.[6] The method minimizes the objective function by iteration its evaluation at each vertex and by replacing point below as : Order, Reflect, Expand, Contract, Shrink.

III. MODEL BUILDING AND TESTING

A. Vehicles Types Dataset Definition

Most of the cars in Myanmar are modified used vehicles from Asia Car Makers, so the accuracy of famous datasets had questioned as Stanford cars dataset, which includes most US car makers. [7] Besides, the class of reference Stanford Cars Dataset has Car Maker, model and Years, the dataset modified the Vehicle types in 7 classes with relevant environment – Motorbikes (0), Saloon (1), Wagon (2), Van (3), SUV (4), Truck (5) and Bus (6). [8] This research created new Myanmar Cars datasets, collected and organized vehicles image from Facebook used cars auctions page as collection of 5043 images. Car maker, model and year had not considered in this system.

B. System Implementation and model training

All Image Datasets are trained with openCV2, Caffe model building to get training image set and added hyper-parameters optimization the objective function by repeating its evaluation vertex. The system run on AWS instance with 8 core vCPUs, 16 GB RAM and powerful NVIDIA GPU with 1,536 cores (4GB of video RAM). In training process, all images are divided into 5 parts 4 parts is used train image sets and the remaining 1 for test sets. Both sets have .csv files with training and test images labels. To establish and generate an pre-defined parameters randomly, then complete the training model run up to 20000 iterations.

IV. EXPERIMENTAL RESULT

This system achieved that the classification and accuracy of detection vehicles for wedding ceremony video streaming, loss and accuracy for learning rate is better accurate in Myanmar Cars Dataset trained system. Motorbikes are not allowed in study area; hence it cannot count in the result.

TABLE II. AVERAGE ACCURACY OF DIFFERENT TYPES OF VEHICLES

Types	Average Accuracy (%)	
	Myanmar Car Dataset	*Stanford Car Dataset*
Saloon	84.5%	81.4%
Van	77.3%	73.1%
Wagon	69.5%	66.50%
Bus	82.6%	76.4%
Motorbikes	0%	0%
Truck	78.3%	64.2%
SUV	63.0%	59.5%

V. CONCLUSION AND FUTURE WORKS

The vehicles classification and counting system in this paper presented that the result had much better by using deep learning image classification with hyper-parameter optimization and new relevant environment Dataset. By incorporating hyper-parameters, the model is more robust to handle the large dataset of vehicles. The major contributions are (1) preparing Myanmar Cars Image Dataset and modifying reference Stanford Cars dataset (2) using the hyper-parameters to deep learning model applied in Myanmar events. The results perform well and innovative in deep neural learning with additional parameters and datasets. To cover more variations of Myanmar vehicles, Cars Image datasets need more data to classify and train.

REFERENCES

[1] Han Y, Jiang T, Ma Y, Xu C. "Pretraining convolutional neural networks for image-based vehicle classification". Advances in Multimedia. 2018 Jan 1;2018.

[2] Song, Huansheng, et al. "Vision-based vehicle detection and counting system using deep learning in highway scenes." European Transport Research Review 11.1 (2019): 51.

[3] Fan, Q., Brown, L., Smith, J. (2016). "A closer look at faster r-cnn for vehicle detection", In 2016 IEEE intelligent vehicles symposium (IV). https://doi.org/10.1109/ivs.2016.7535375: IEEE.

[4] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '14), pp. 580–587, Columbus, Ohio, USA, 2014.

[5] R. Girshick, "Fast R-CNN," in Proceedings of the 15th IEEE International Conference on Computer Vision (ICCV '15), pp.1440–1448, Santiago, Chile, 2015

[6] Ozaki, Yoshihiko, Masaki Yano, and Masaki Onishi. "Effective hyperparameter optimization using Nelder-Mead method in deep learning." *IPSJ Transactions on Computer Vision and Applications* 9.1 (2017): 20.

[7] Krause J, Stark M, Deng J, Fei-Fei L. "3d object representations for fine-grained categorization". InProceedings of the IEEE international conference on computer vision workshops 2013 (pp. 554-561)

[8] Automotive in Myanmar, IPSOS Business Consulting, 2013, https://www.ipsosconsulting.com/insights/publications/automotive-in-myanmar

978-1-7281-9802-6/20/$31.00 ©2020 IEEE