QS STARS RATED FOR EXCELLENCE 2012

Fpt University

TRƯỜNG ĐẠI HỌC FPT

Improving Warped Planar Object Detection Network For Automatic License Plate Recognition

Member:

Nguyen Dinh Tra, Nguyen Cong Tri, Vu Thanh Trung

Advisor

Assoc. Prof. Duy Hung Phan

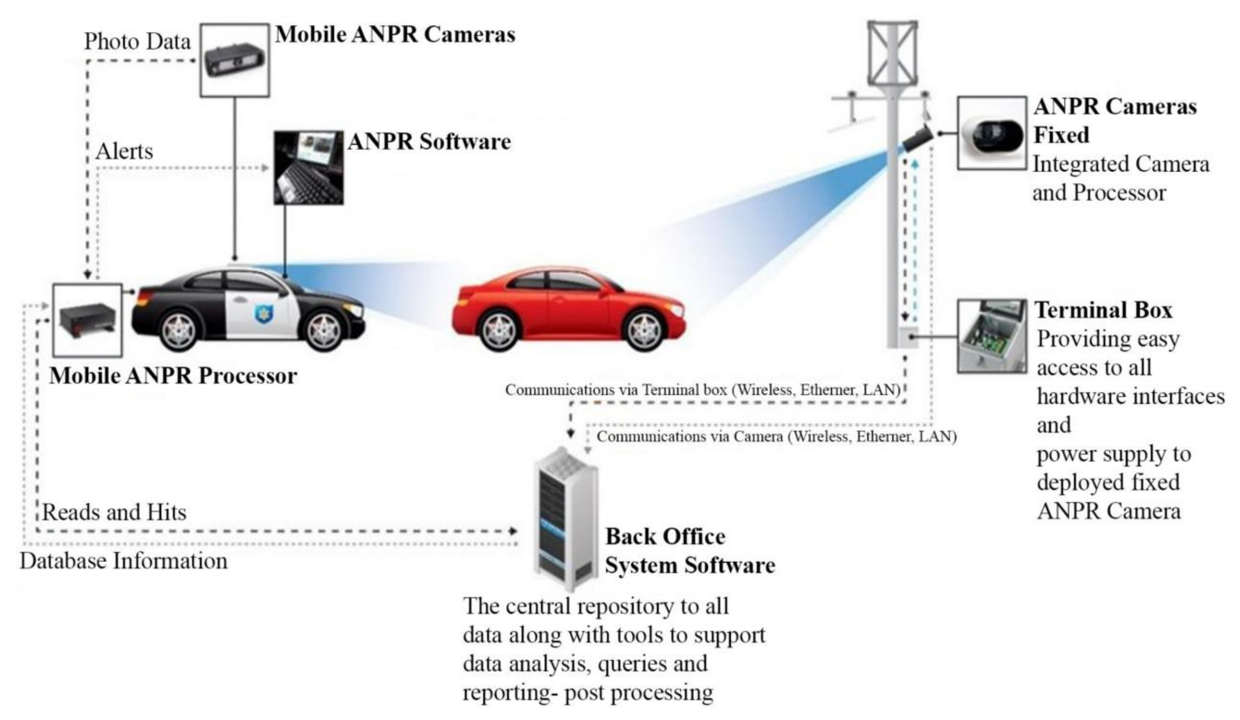
FPT University, HaNoi, VietNam

Table of Contents

01 Introduction

Overview

- The problem of recognizing license plates is not a new problem. The application of license plate recognition is widely used in automatic ticketing parking lots, toll stations on highways or as vehicle tracking systems in traffic and detecting vehicles
- Currently, there are many different methods proposed to solve this problem such as YOLO, SSD, STN.

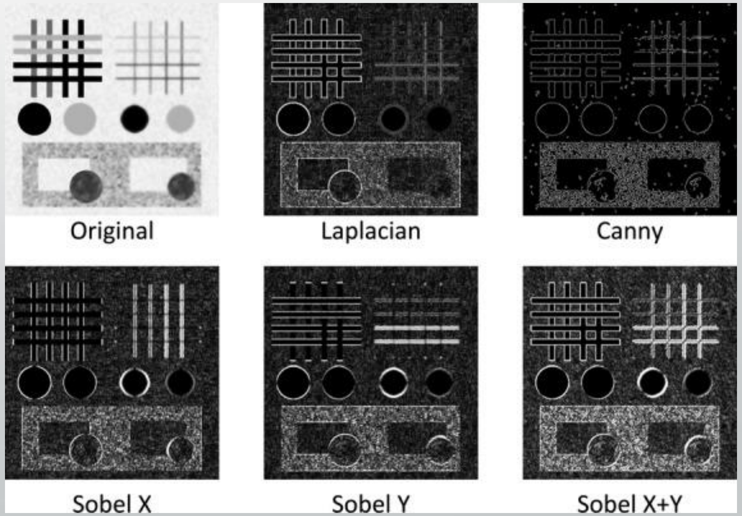


Main Problem

Related works

- Sergio et al. designed WPOD-Net based on the idea of YOLO, SSD, STN (Spatial Transformer Network). As mentioned above YOLO and SSD only return 1 bounding box in the license plate regardless of the surrounding space. STN can detect non-rectangular regions, but it cannot handle multiple transformations at the same time, instead performing a single spatial transformation across the entire input. WPOD-Net return bounding box area surrounds the license plate and brings the number plate to the front view.

- In another paper published in 2009: “A Descriptive Algorithm for Sobel Image Detection” , Vincent et al. use a Gaussian filter to remove noise, smooth the image first to make the edge detection algorithm work better.



Contribution

- Our goal in this thesis is to improved Warped Planar Object Detection(WPOD-Net) in end-to-end automatic license plate recognition using feature engineering to increase accuracy.
- This study suggests a methods when building model in Computer Vision, namely the use of edge detection as a layer to enhance the learning efficiency of the model. And using qIOU metric helps find overlap between ground truths bounding box and predicted bounding box increasing the accuracy more than using IOU metrics in distorted object.

02
Background

Complete ALPR Methods

Pipeline of Automatic License Plate Recognition

Edge Detection

- Sobel Edge Detection

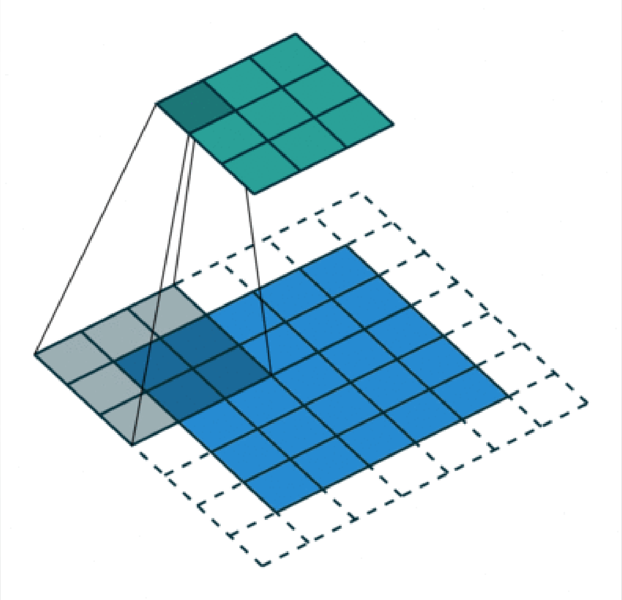
-1	0	+1
-2	0	+2
-1	0	+1

Gx

+1	+2	+1
0	0	0
-1	-2	-1

Gy

3x3 Sobel edge detection masks

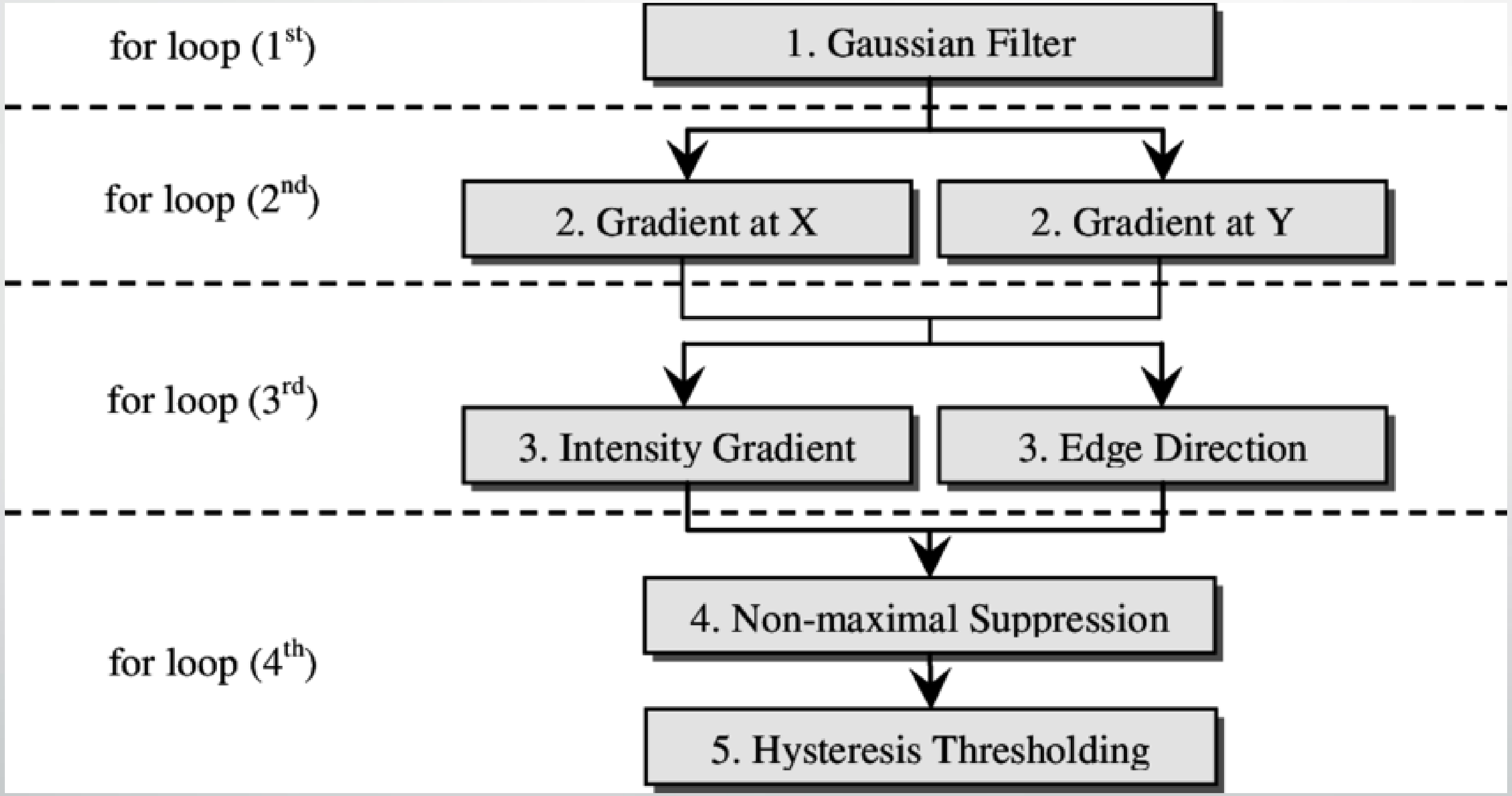


Visualization convolution kernel

Edge Detection

Edge Detection

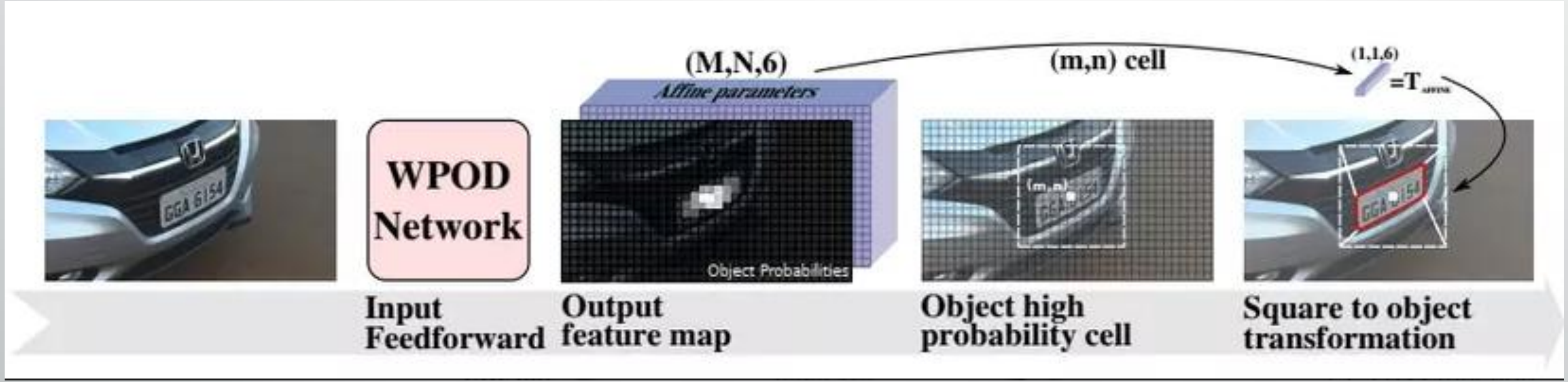
- Canny edge detector



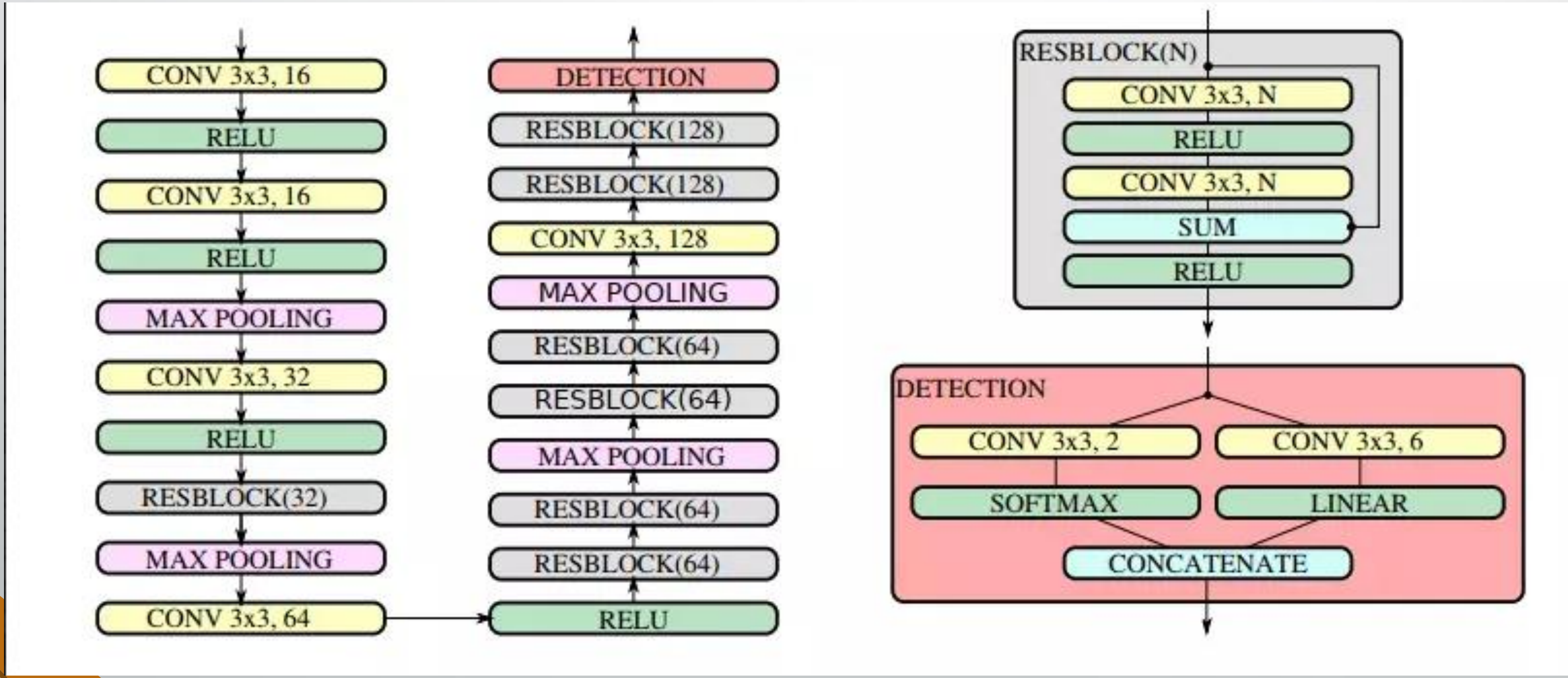
Edge Detection

03
Methodology

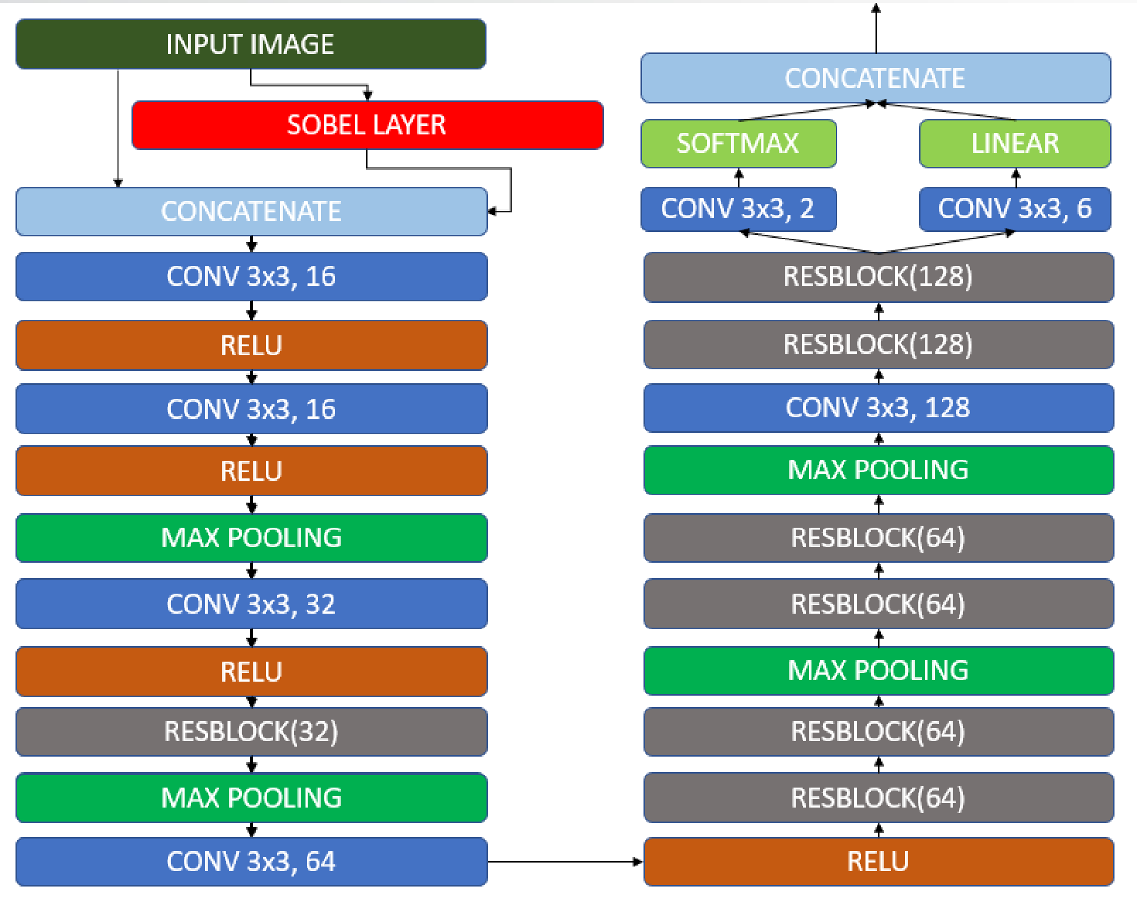
How the WPOD-Net work?

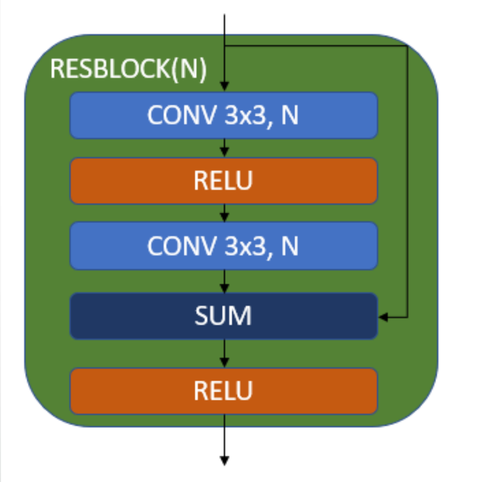


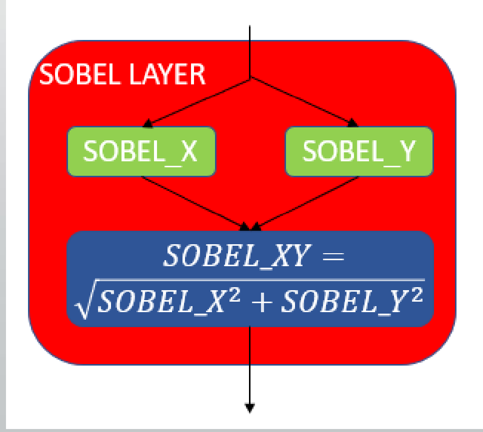
The Architecture WPOD-Net



Propose Model Architecture



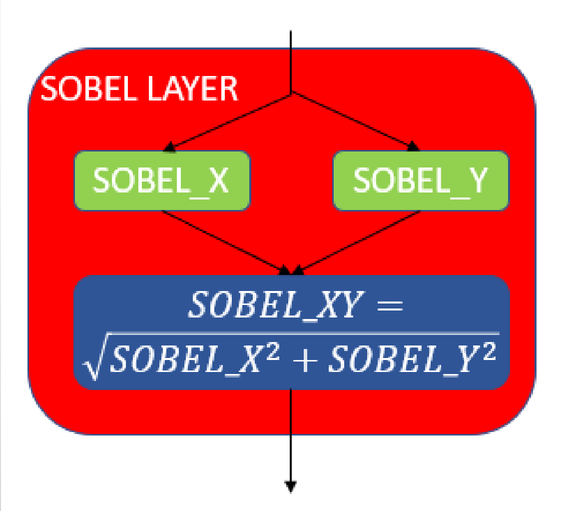




Sobel layer

We consider $\mathbf{G}_x$ and $\mathbf{G}_y$

$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \text{ and } \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$



Combining the gradient

$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}$$

Loss Function

Local affine transformation

$$T_{mn}(\boldsymbol{q}) = \begin{bmatrix} \mathrm{m}\ \ (v_3, 0) & v_4 \\ v_5 & \mathrm{m}\ \ (v_6, 0) \end{bmatrix} \boldsymbol{q} + \begin{bmatrix} v_7 \\ v_8 \end{bmatrix},$$

Set v_{i} is the i-th dimension on the output features map, (m,n) is the coordinates of the point being considered on the output features map.

Normalization function

$$A_{mn}(\boldsymbol{p}) = \frac{1}{\alpha}\left(\frac{1}{N_s}\boldsymbol{p} - \begin{bmatrix} n \\ m \end{bmatrix}\right),$$

Set $\alpha = 7.75$, $N_s = 4$

Loss Function

- Location loss:

$$f_{location}(m,n) = \sum_{i=1}^{4} (T_{mn}(\boldsymbol{q}_i) - A_{mn}(\boldsymbol{p}_i))^2$$

- Confidence loss:

$$f_{\text{confidence}}(m,n) = \log\text{loss}(\mathbb{I}_{obj}, v_1) + \log\text{loss}(1 - \mathbb{I}_{obj}, v_2)$$

- Total loss:

$$\text{Total\_loss} = \sum_{m=1}^{M} \sum_{n=1}^{N} [\mathbb{I}_{obj} f_{location}(m,n) + f_{\text{confidence}}(m,n)]$$

$\mathbb{I}_{obj}$ is the object indicator function that return 1 if there is an object at point (m,n) or 0

04
Experiments and Results

Training Dataset

- Training Dataset:363 images take AOLP dataset, Vietnamese car and motorbike license plates we collected from Internet.

AOLP dataset: 50 images

Cars: 105 images

Motorbike: 208 images

- Augmentation transforms technique: Rectification, Aspect-ratio, Centering, Scaling, Rotation, Mirroring, Translation, Cropping, Color space, Annotation.
- Test Dataset: 175 images include:

AOLP dataset: 25 images

Cars: 50 images

Motorbike: 100 images

Training Dataset

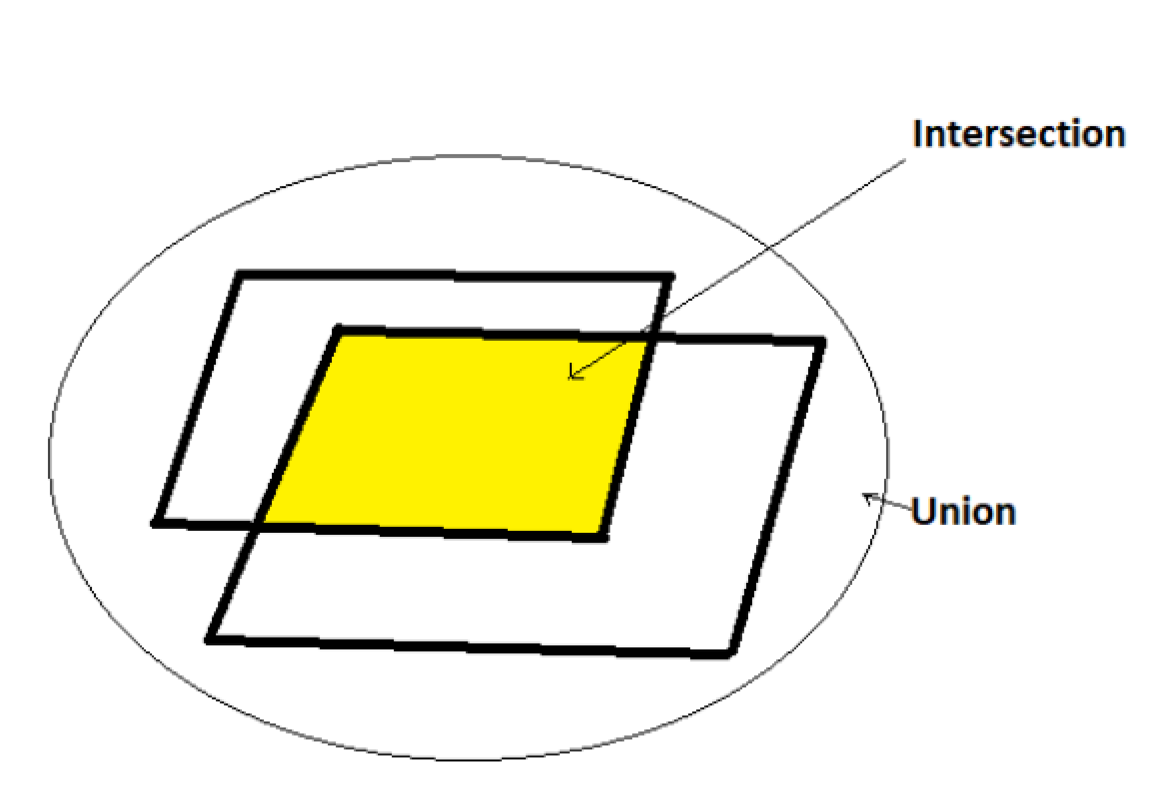
Training Details

- Training time: 25 hours (1 GPU 3090 RTX)
- GPU memory usage: 24GB
- Batch size: 64
- Learning rate: 10^{-3}
- Iterations: 300000

Evaluation metrics

- qIOU(Quadrilateral Intersection over Union) metric for plate detection

$$qIOU = \frac{Intersection}{Union}$$



Evaluation metrics

- Accuracy

$$\text{accuracy} = \frac{TN+TP}{TN+TP+FP+FN}$$

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

Evaluate Dataset

- AOLP dataset

Road patrol (RP)

- Our test-set: 175 images include:

AOLP dataset: 25 images

Cars: 50 images

Motorbike: 100 images

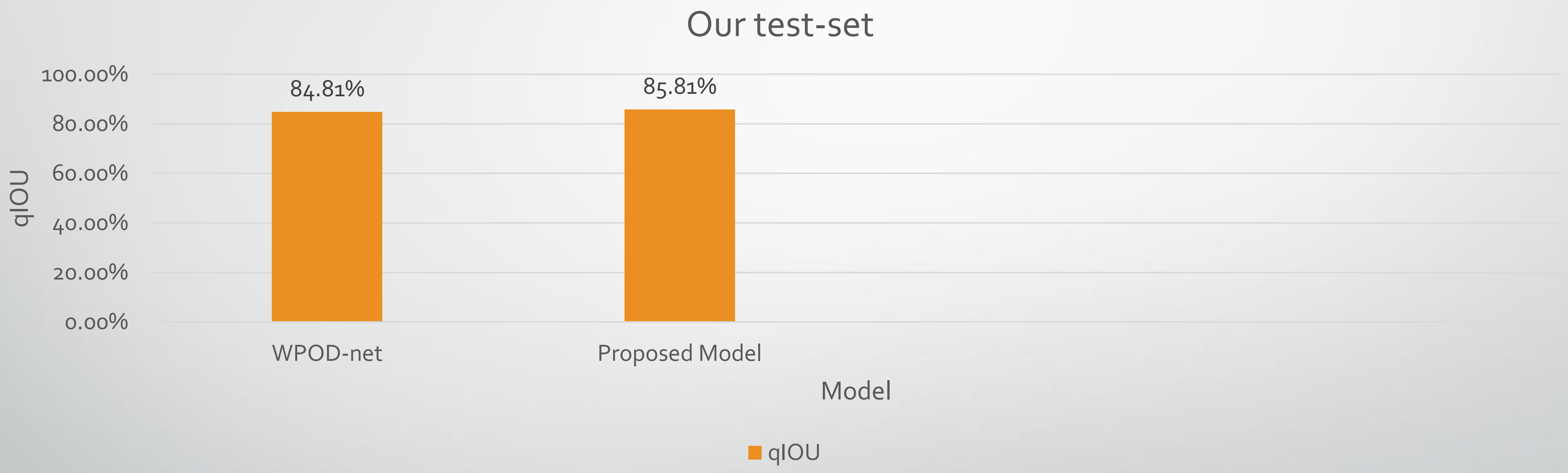
Evaluate Dataset

- OpenALPR dataset:

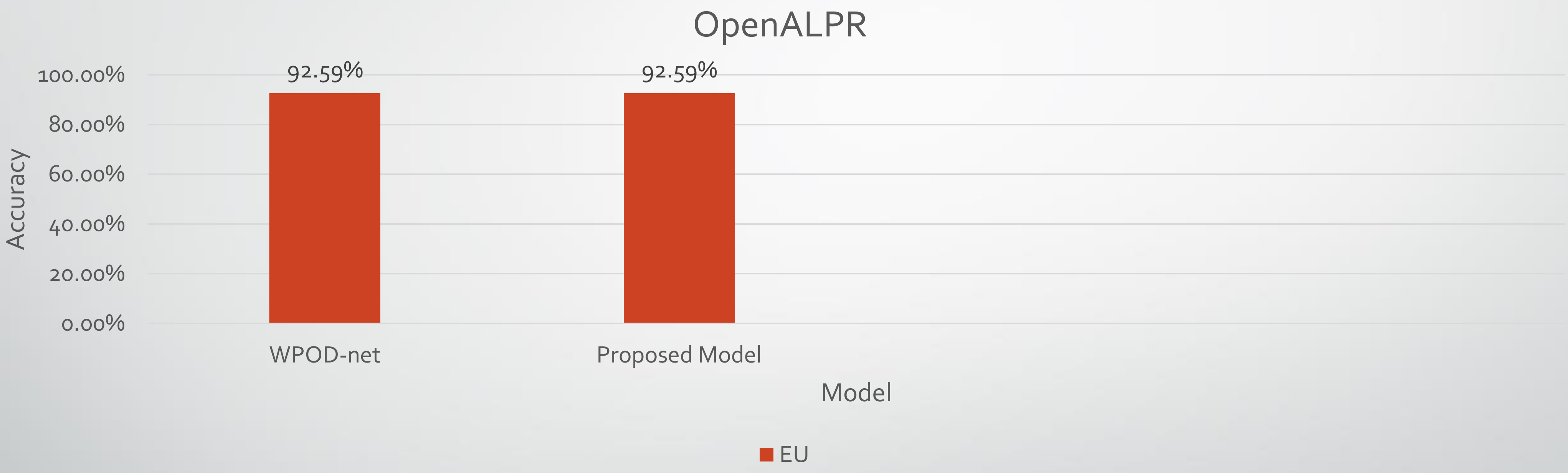
  EU license plates

  Brazil license plates

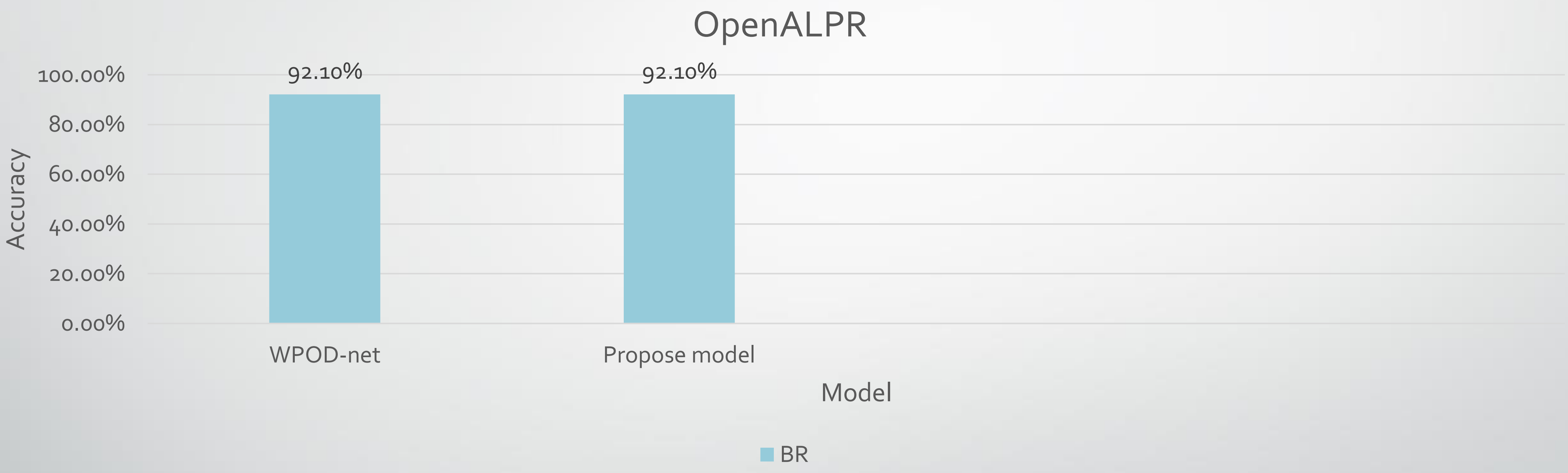
Comparison Between the proposed model and WPOD-Net in our test dataset



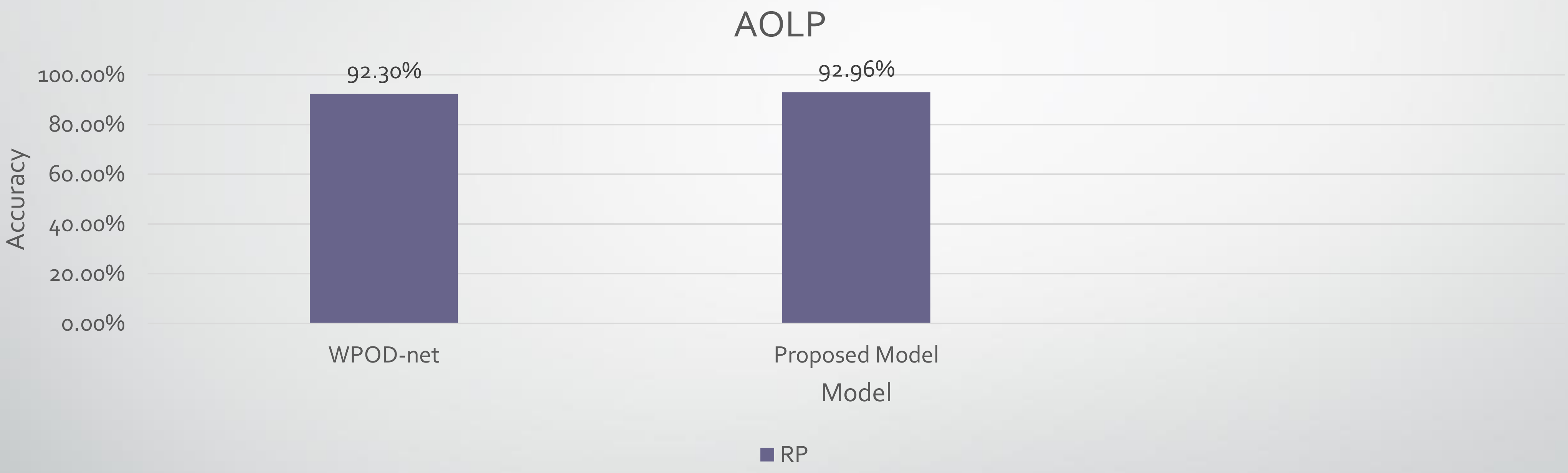
Comparison Between the proposed model and WPOD-Net in OpenALPR-EU (in Complete ALPR method)



Comparison Between the proposed model and WPOD-Net in OpenALPR-BR (in Complete ALPR method)



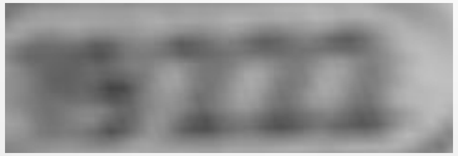
Comparison Between the proposed model and WPOD-Net in AOLP (in Complete ALPR method)



Result


ExoticsOnRoad.com
HH P 8586


HH P 8586

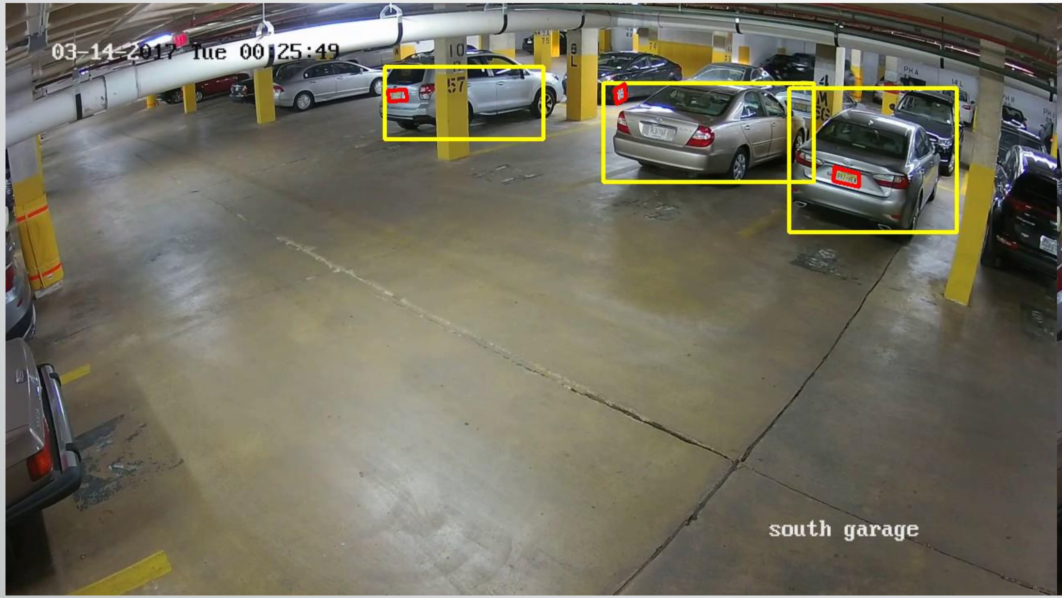



HH P 8586


ExoticsOnRoad.com
HHP8586

Qualitative Comparison

WPOD-Net	Proposed Model



Qualitative Comparison

WPOD-Net	Proposed Model
	

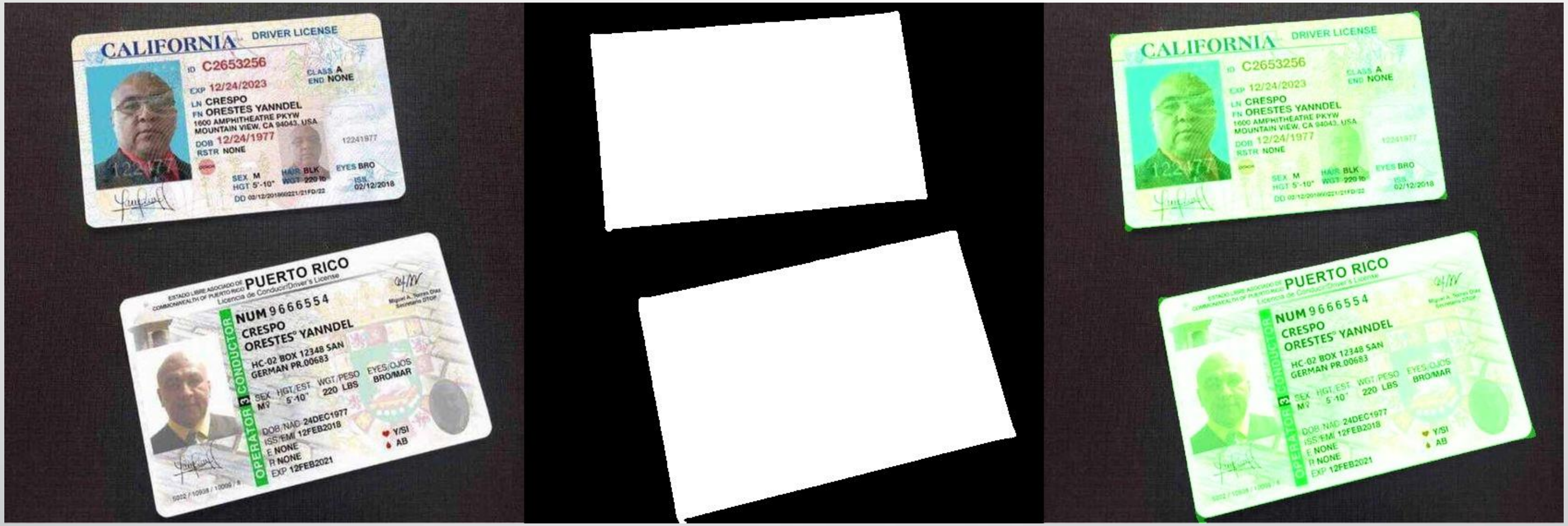
05
Conclusion and future work

Conclusion

- Improve the performance of WPOD network.
- Accurate object detection, especially in cases where the subject imaged is tilted.
- The thesis can be a good reference for many machines learning and data mining problem.

Future Work

- Applications of this model to other practical problem such as menu digitization or document digitization.



- Robust network by changing the backbone of the network or using another feature extraction methods.

The End

Question?