

License Plate Detection using Deep Learning and Font Evaluation

Rabiah A. Al-qudah

A Thesis

In

The Department

of

Computer Science and Software Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science (Computer Science) at

Concordia University

Montreal, Quebec, Canada

January 2019

© Rabiah Al-qudah, 2019



## ABSTRACT

### License Plate Detection using Deep Learning and Font Evaluation

Rabiah Al-qudah

License plate detection (LPD) in context is a challenging problem due to its sensitivity to environmental factors. Moreover, the chosen font type in the license plate (LP) plays a vital role in the recognition phase in computer-based studies. This work is two folded. On one hand, we propose to employ Deep Learning technique (namely, You Only Look Once (YOLO)) in the LPD. On the other hand, we propose to evaluate font characteristics in the LP context.

This work uses 2 different datasets: UFPR-ALPR, and the newly created CENPARMI datasets. We propose a YOLO-based adaptive algorithm with tuned parameters to enhance its performance. In addition to report the recall ratio results, this work will conduct a detailed error analysis to provide some insights into the type of false positives. The proposed model achieved competitive recall ratio of 98.38% with a single YOLO network.

Some fonts are challenging for humans to read; however, other fonts are challenging for computer systems to recognize. Here, we present 2 sets of results for font evaluation: font anatomy results, and commercial products recognition results. For anatomy results, 2 fonts are considered: Mandatory, and Driver Gothic. Moreover, we evaluate the effect of the used fonts in context for the two datasets using 2 commercial products: OpenALPR and Plate Recognizer. The font anatomy results revealed some important confusion cases and some quality features of both fonts. The obtained results show that the Driver font has no severe confusion cases in contrast to the Mandatory font.

## **ACKNOWLEDGEMENTS**

My sincere gratitude goes to my supervisor Prof. C.Y.Suen. I would not have been able to finish this work without his valuable advice, assistance, and encouragement. I am very grateful for the outstanding supervision I have received.

I would like to thank the examination committee members, Drs. T. D. Bui, and T. Fevens for their efforts and valuable feedback.

## **DEDEICATION**

I am deeply indebted to my parents, Dr Abdulhamid and Mrs Nedal Al-qudah for their continuous encouragement, support and patience during the years I was away from home. My deep appreciation goes to my brothers and sisters who were all been a source of support.

Finally, my deepest gratitude and appreciation goes to my small family; to my dear husband Dr Yaser Khamayseh, for his love, support, and patience which has enabled me to finish my graduate study. To my daughters, Yara and Hala for the love and joy they have added to my life, and for always encouraging me to work even when I knew that, at their young age, they wanted me to be with them.

## Table of Contents

Chapter 1. Introduction .....	1
I. License Plates.....	1
II. License Plate Detection.....	4
III. Fonts.....	5
IV. Digital Font Evaluation .....	7
V. Thesis Structure .....	8
Chapter 2. Literature Review .....	9
I. Machine Learning Based Techniques.....	9
II. Deep Learning Based Techniques .....	12
III. Font Evaluation.....	15
Chapter 3. Proposed Model.....	19
I. Datasets .....	19
II. License Plate Detection.....	24
III. YOLO.....	24
IV. Proposed Detection Methodology.....	25
V. Digital Font Evaluation .....	29
Chapter 4. LPD Results and Analysis.....	31
I. LDP Results.....	31
II. Error Analysis .....	37
Chapter 5. Font Evaluation Results and Analysis.....	44
Chapter 6. Conclusion and Future Work .....	56
I. Conclusion.....	56
II. Future Work.....	57

## List of Figures

Figure 1. Sample Brazilian License Plate .....	3
Figure 2. Sample of Quebec License Plate .....	3
Figure 3. Example of Font Characteristics [7].....	7
Figure 4. Sample Images from UFPR-ALPR Dataset .....	20
Figure 5. Letter and Digit Distributions in UFPR-ALPR.....	21
Figure 6. Sample Images from CENPARMI Dataset .....	23
Figure 7. Letter and Digit Distributions in CENPARMI Dataset.....	24
Figure 8. Proposed LP Detection Algorithm .....	26
Figure 9. Recall Results for YOLO Models Trained with 2700 Instances.....	31
Figure 10. Recall Results for YOLO Models Trained with 1800 Instances.....	32
Figure 11. Sample Detection Results for the YOLO Network Trained with 800 Resolution and Tested with 960 Resolution from CENPARMI Dataset.....	35
Figure 12. Sample Detection Results for the YOLO Network Trained with 800 Resolution and Tested with 960 Resolution from UFPR-ALPR Dataset.....	36
Figure 13. Number of Iterations for Different Training Models .....	37
Figure 14. Distribution of False Positives at Training Network Input Resolution 320 and Dataset Size 2700 .....	38
Figure 15. Distribution of False Positives at Training Network Input Resolution 608 and Dataset Size 2700 .....	38
Figure 16. Distribution of False Positives at Training Network Input Resolution 800 and Dataset Size 2700 .....	39
Figure 17. Distribution of False Positives at Training Network Input Resolution 960 and Dataset Size 2700 .....	39
Figure 18. Distribution of false positives at training network input resolution 320 and dataset size 1800.....	40
Figure 19. Distribution of False Positives at Training Network Input Resolution 608 and Dataset Size 1800 .....	40
Figure 20. Distribution of False Positives at Training Network Input Resolution 800 and Dataset Size 1800 .....	41
Figure 21. Distribution of False Positives at Training Network Input Resolution 960 and Dataset Size 1800 .....	41
Figure 22. Sample Image for Detecting Wrong Object .....	42
Figure 23. Sample Image for a LP Considered as FP Caused by Dataset Lack of Annotation .....	42
Figure 24. Sample Image for Error Caused by $0 < \text{IoU} < \text{Threshold}$ .....	43
Figure 25. Mandatory Font Glyphs.....	44
Figure 26. Detection Distribution for all Letters and Digits using Plate Recognizer for UFPR-ALPR Dataset .....	48
Figure 27. Heat Map of the Confusion Matrix for Plate Recognizer using UFPR-ALPR Dataset .....	48
Figure 28. Detection Distribution for all Letters and Digits using OpenALPR for UFPR- ALPR Dataset.....	49

Figure 29. Heat Map of the Confusion Matrix for OpenALPR using UFPR-ALPR Dataset...	50
Figure 30. Detection Distribution for all Letters and Digits using OpenALPR for CENPARMI Dataset .....	52
Figure 31. Heat Map of the Confusion Matrix for OpenALPR using CENPARMI Dataset....	52
Figure 32. Detection Distribution for all Letters and Digits using Plate Recognizer for CENPARMI Dataset .....	53
Figure 33. Heat Map of the Confusion Matrix for Plate Recognizer using CENPARMI Dataset .....	53

## List of Tables

Table 1. Font Characteristics .....	6
Table 2. Distribution of Letters and Digits in UFPR-ALPR Dataset .....	22
Table 3. Training YOLO network structure .....	27
Table 4. Testing YOLO network structure .....	28
Table 5. Font Type Evaluation Aspects.....	30
Table 6. Comparison of Recall Ratio Results.....	37
Table 7. Mandatory Font Anatomy.....	45
Table 8. Quebec Font Anatomy .....	46
Table 9. Recall Ratio Results.....	47
Table 10. UFPR-ALPR Dataset Classification Percentages using both Products .....	51
Table 11. CENPARMI Dataset Classification Percentages using both Products.....	54

# Chapter 1. Introduction

License plate detection (LPD) in context is a challenging problem due to its sensitivity to environmental factors (such as rain, dust, and shadow) and light, which may greatly influence the detection accuracy. Moreover, LPD is more challenging for real time systems. Furthermore, the chosen font type in the License Plate (LP) plays a vital role in the recognition phase in computer-based studies. Despite its vital importance and its great effect on the recognition accuracy, font evaluation for recognition is not sufficiently investigated in the literature. Hence, this work is two folded. On one hand, we propose to employ Deep Learning (DL) techniques in the LPD. On the other hand, we propose to evaluate font characteristics in the LP context. The following sections will provide more details about license plates, license plate detection, fonts, and font evaluation.

## I. License Plates

License Plate, also referred to as number plate or vehicle registration plate, is usually a rectangular shaped metal/plastic object attached to the motor vehicle at the back, front or both side(s) for official identification. It is a mandatory requirement for vehicles to be legally allowed to be driven on public roads. Each jurisdiction has an official agency that oversees the registration and issuing of LP, in which a unique identification number is assigned for each vehicle/plate. Plate number consists of letters or digits or a combination of both.

Each jurisdiction has its own regulations regarding the design and the placement of LP. Some jurisdictions require the LP to be attached to the front and the rear of the vehicle, other jurisdictions require the LP to be attached to the vehicle's rear only.

The importance of LPs is apparent in emergency cases such as accidents and car thefts, traffic violations, and illegal usage of vehicles. LP helps to identify the vehicle and its owner in a timely and efficient manner. LP importance is even more apparent with the rise of automatic detection systems. To improve the detection and recognition of LPs, each jurisdiction defines and enforces its own design for LPs. The LP design defines the following aspects:

- Typeface, nowadays, most jurisdictions use anti-fraud typefaces (such as FE-Schrift typeface) to prevent fraudsters from altering numbers or characters to resemble other characters.

- Size, the plate size (such as, width and height). The LP size must be adequate to fit the numbers and letters inside the license plate without being excessively tight or excessively loose. In fact, there are 3 basic standards adopted worldwide for LP size [8]:
  - 520 mm by 110 mm used by most European countries.
  - 305 mm by 152 mm used in North America and Central America, and parts of South America.
  - 372 mm by 135 mm used in Australia.

However, some jurisdictions may adopt their own standard.

- Decoration (if any). Some jurisdictions (or territories) may add certain decoration items to the LP such as national symbol, slogans, and/or background pictures.
- Combination of letters and numbers. The LP registration number consists of either letters or digits or a combination of both. The length of the LP number determines the number of unique combinations, and hence, the total number of LP that can be issued. For example, if a LP is made of 5 numbers and 1 letter, then the number of unique combinations is 2,600,000.

Again, the choice of the typeface is very important in the recognition process, typically, typefaces were chosen to be perceived swiftly and precisely under daylight conditions from a reasonable distance by human (naked eyes).

This work considers 2 datasets, the first dataset is from Brazil, while the second dataset is from Quebec/Canada. The following subsections will provide more details on the LP used in Brazil and Quebec.

### **1. Brazilian License Plates**

LP in Brazil uses the ABC.1234 form. Once a combination is assigned to a vehicle it can not be transferred to another one. A metallic band with the state abbreviation (e.g., RJ = Rio de Janeiro) and the municipality name is placed on top of the combination. Figure 1 depicts a Brazilian LP.



Figure 1. Sample Brazilian License Plate

LPs have different sizes for cars and motorcycles: 400 mm x 130 mm is used for cars, and 200 mm x 170 mm is used for motorcycles. Moreover, LPs have different colors based on the category of the vehicle. For example, LP for private vehicles are gray, while public transportation vehicles, such as buses and taxis have red LPs. The typeface used in LPs is Mandatory.

As of 2018, Brazil will adopt a new LP design, and by 2023, all vehicles will be equipped with the new plates. However, the old LP design is used by all vehicles in the dataset.

## 2. Quebec License Plates

Over the years, many designs were adopted in Quebec for LPs. Moreover, there are many categories of LPs based on vehicle's type and use. Recently, the province of Quebec allowed the personalized LPs (also called vanity plates). Since 2010, the private passenger cars used the A12-BCD format with the slogan "Je me souviens" centred at the bottom, and the word "Quebec" at the top. Figure 2 depicts a sample of recent LP from Quebec for commercial vehicle.



Figure 2. Sample of Quebec License Plate

In Quebec, there are 15 categories of non-personalised licence plates assigned based on vehicle type and use. Currently, letters I, O, and U are not used in non-personalised LPs [2]. Personalised LPs are permitted in Quebec, and must obey the following rules:

1. Spaces are not permitted.
2. The licence plate character combination must not be easily confused with another licence plate number.

3. The character combination must be clear and easy to read.
4. Combinations with more than 4 consecutive identical characters are not permitted (e.g.: AAAAA).
5. The personalized character combination chosen must not include an expression or message that, when read forwards or backwards:
  - a. Gives the false impression that the road vehicle's owner is, or is connected to, a public authority
  - b. Expresses disregard toward road safety
  - c. Expresses an obscene, scandalous or sexual idea
  - d. Expresses abusive or derogatory language or offensive expressions in any language, including texting language
  - e. Promotes the perpetration of a criminal offence
  - f. Is restricted or prohibited by law (e.g. trademark). It must not violate any intellectual property rights.

## **II. License Plate Detection**

The usage of deep learning attracted the attention of researchers in recent years. It is being widely employed to solve classification and detection problems. It is used in many domains such as object detection and recognition, natural language processing, and computer aided diagnosis. For object detection and recognition, many techniques have been developed such as Convolutional Neural Network (CNN). However, in the past 2 years, a new model was developed, namely You Only Look Once (YOLO) [1]. YOLO is a real time object detection system that performs objects' detection and objects' classification in one scan rather than performing these 2 operations in 2 scans.

The success of deep learning algorithms depends heavily on the available training dataset. This work used Federal University of Paraná dataset for Automatic License Plate Recognition (UFPR-ALPR) which consists of 4500 images of LPs from Paraná /Brazil. This dataset was captured with three different cameras; GoPro Hero4 Silver, Huawei P9 Lite and iPhone 7 Plus. It consists of images taken from motorcycles, cars and public transportation vehicles. The diversity of image brightness setting, vehicle distance, LP sizes and LP background colors are challenging factors for LPD.

Since YOLO is a new technique, the choice of certain parameters can affect the detection accuracy. Hence, we propose a YOLO-based adaptive algorithm that tunes its parameters to enhance its performance. The following experiments are conducted:

1. Training and testing network resolution effect.
2. Dataset size.

Moreover, we propose to train the YOLO net using pre-trained weights on Common Object in Context (COCO) dataset instead of randomly initialized weights [2]. To improve the detection process, we propose to train the model using one single YOLO net instead of two nets.

To evaluate the obtained results from the above-mentioned experiments, we propose to use the following metrics:

1. Recall ratio: the ratio of correctly predicted objects to the total number of relevant objects. Mathematically, it is the ratio of true positives (TP) to the summation of TP and false negatives (FN). TP is the number of correctly identified objects. And, FN is the number of misdetected objects.
2. Error analysis, it will provide some insights into the type of false positives (FP) [3]. FP is the number of faulty detected objects.

### **III. Fonts**

Font is a collection of printable and displayable text characters that is defined by specific style and size. A set of fonts that share common design features is called typeface. Each font is defined by specific style, condensation, weight, width, italicization, slant, ornamentation, and designer or foundry. Typeface is a set of glyphs that represent an individual letter, number, punctuation mark, or symbol. Typefaces can be designed for special applications such as mathematics, and LPs.

The choice of font or typeface plays a vital role in the recognition process of LPs. Some fonts are challenging for humans to comprehend; however, other fonts are challenging for computer systems to recognize.

Different fonts exhibit various degrees of readability and emphasis. In fact, some fonts are designed to make them of more visual interest. Each font is distinguished by a unique combination of characteristics. Font's characteristics can affect the font readability, legibility, and

recognizability by both humans and computer systems. Table 1 lists some of the most important characteristics and their definition [6]. Figure 3 shows an example of these features and how they relate to sample letters. The figure is extracted from [7].

Table 1. Font Characteristics

Characteristics	Definition
Weight	Thickness of the character outlines relative to their height
Slope	Used to represent italic type or oblique type to emphasise important words
Width	Represents the character's width or stretch
Serif and sans serif	A small line attached to the end of a stroke in characters. Fonts can be either serif or sans serif
Baseline	The optical line on which type sits
Mean line	An optical line created by the eye moving across the top of a set of lowercase letters
Ascender	The stem of a lowercase letter which extends past the mean line
Descender	The stem of a lowercase letter which extends below the baseline
x-height	The height of the lowercase letters without descenders or ascenders
Cap line	An optical line created by the eye moving across the top of a set of uppercase letters
Bowl	Fully or modified rounded forms found in letters such as C
Counter	Negative space fully or partially enclosed by the stroke of a letter form like in B
Tail	Short downward strokes
Crossbar	Horizontal stroke connecting two parts of a letterform like in H
Apex	Juncture of a stem like in letters A and M
Spur	A projection smaller than a serif, that reinforces the point at the end of curved stroke. It can be found in G

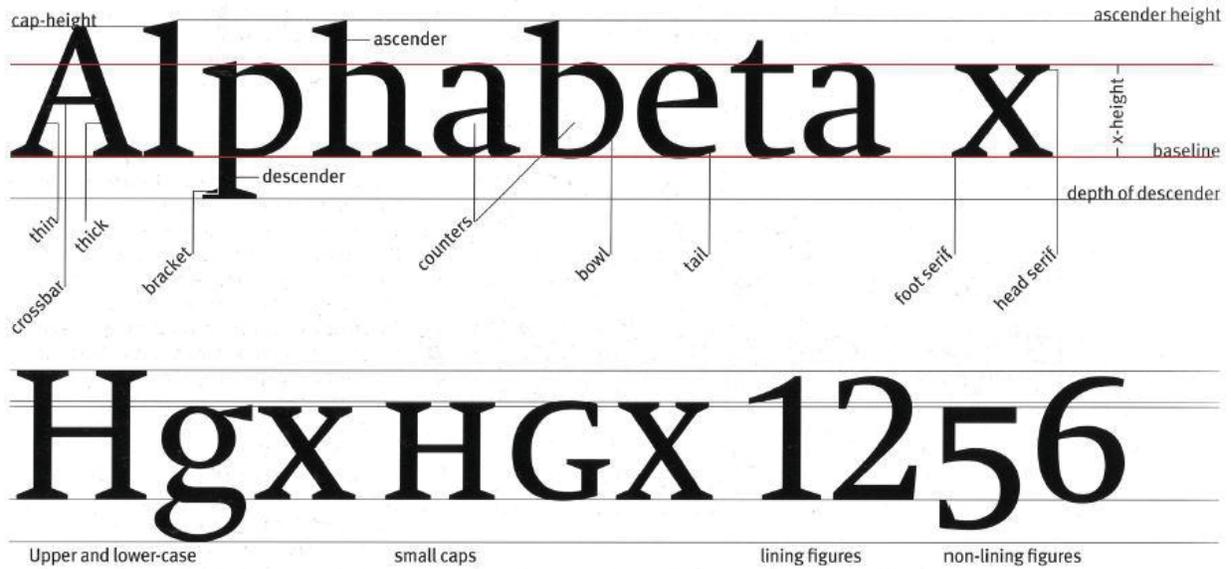


Figure 3. Example of Font Characteristics [7]

In computer systems, choosing the right font plays a crucial role in the system's ability to recognize the underlying text.

#### IV. Digital Font Evaluation

Automatic License Plate Recognition (ALPR) consists of 3 main steps: detection, segmentation, and recognition. We note that even the best segmentation algorithms do not always produce recognizable letters and digits from the cropped segments, hence, each character must maintain unique integrity to overcome cropping shortcomings. Therefore, choosing the adequate font is a key to successful LP recognition, even for general plate numbers that do not support regular expressions.

This work proposes the following contributions:

1. Study the characteristics of 2 different fonts: Mandatory, and the font used for license plates in Quebec. Fourteen different font cases will be considered. The cases are: A and 4, B and 8, C and G, D and P, E and F, I and 1, 8 and 6, 0 and O, Q and O, U and V, V and Y, Z and 2, S and 5, G and 6 designs.
2. Evaluate the effect of the used fonts in context for 2 datasets: UFPR-ALPR dataset for the Mandatory font, and Quebec LPs dataset for Quebec LP font.

The evaluation will be conducted using 2 commercial products: OpenALPR [4], and Plate Recognizer [5]. OpenALPR is an automatic number-plate recognition system that is developed by OpenALPR Technology Inc. It was released in May 2014 and can detect LP from many countries such as Canada, Brazil, USA, Japan and China. Plate Recognizer is an automatic license plate recognition API. Its training dataset is collected from more than 100 countries and regions around the world. It uses deep neural networks for the core detection module. It uses two steps to perform the recognition from image to plate numbers. In the first step, it detects the license plate, and in the second step it produces each characters in the detected license plate.

Based on the obtained results, we suggest a set of recommendations for some adequate font characteristics to enhance recognition accuracy in real time.

## **V. Thesis Structure**

The rest of this thesis is organized as follows: Chapter 2 provides a review on the literature work for both LP detection and font evaluation. Chapter 3 introduces the proposed system model. Chapter 4 presents and discusses the obtained results. Finally, Chapter 5 concludes this work and suggests potential research direction for future work.

## Chapter 2. Literature Review

Vehicle license plate detection is one of the most widely investigated problems in automatic transportation systems and it attracts the attention of many researchers [9]. It is an active topic in the research community due to its importance in building safe and smart cities. There are many techniques to achieve high detection ratio. Here, we present some of these techniques and discuss their limitations. We divide this part into two sections: Section I presents machine learning-based license plate detection techniques, and Section II presents deep learning-based license plate detection techniques.

For font evaluation, the importance of font choice and its effect on text readability and recognition is discussed and considered by many researchers. This chapter presents some related work for font evaluation in Section III.

### I. Machine Learning Based Techniques

In this section we present some of the techniques that utilize machine learning (ML) for license plate detection.

The work in [10] adopted a cascade detection scheme to detect Chinese license plates; then it proposed a concise and effective license plate verification method. The devised method consists of three key steps: pre-processing, detection and confirmation. In the image pre-processing step, several operations are executed, first, converting RGB image to gray image, then noise elimination (the median filter with  $3 \times 3$  neighborhood), and finally, obtaining gradient image. In the detection of candidate plate regions, cascade AdaBoost classifier by using  $96 \times 26$  training images is used, it is 13 layer cascade AdaBoost classifier with 113 features, which scans across every image by window shifting. Finally, in the candidate verification step, the morphology-based method on the binarized gradient images is used to remove some false positives and to detect plate regions accurately. In the confirmation phase, the off-line trained SVM classifier is employed to confirm the candidate plate regions. 3000 license plate images and 3000 false license plate images are collected to be positive training sample set and negative training sample set respectively. The method in [10] achieved 88.28% recall rate.

The authors in [11] also proposed a three phases solution as follows: pre-processing; sub-image analysis; and license plate classification. The pre-processing phase converts input images to edge images by canny edge detection in 5 steps: images are transformed into greyscale images, and Gaussian filter is used to smooth the result from step 1 to reduce images' noises. For all resulted images, the Gradient's magnitudes and angles are calculated for both vertical and horizontal directions. A non-maxima suppression is applied on the magnitudes and orientations of the gradients. Finally, edges are detected by applying hysteresis thresholding. In sub-image analysis a sliding window technique is used to find Regions of Interest (ROIs). The cropped ROIs are candidates for license plates. The classification phase employs Support Vector Machine (SVM) technique to classify license plates. The proposed SVM uses the linear kernel function.

A Thai license plate dataset is used to evaluate the performance of the proposed method. The proposed method is verified by using 100 images of 100 different cars under different conditions. The proposed method achieved recall ratio of 80%. We note that the used dataset is relatively small to test how generalized the trained model is.

Despite their effectiveness in LP recognition, traditional image processing techniques [12] are restricted by their operations and are limited to certain conditions and environments. Hence, they are not suitable for detecting LP in context. The work in [12] goes beyond the traditional image processing-based techniques. The proposed scheme in [12] performs the feature extraction phase using Deformable Part Models (DPM) from the training images. The extracted features are then fed to Structured SVM (SSVM) classifier. The authors opted to use SSVM instead of SVM classifier because the traditional linear SVM is a binary classifier that is suitable for simple predictions, while the SSVM is a more generalized form of SVM and can achieve a higher prediction accuracy. The detector was trained with only 25-30 annotated images, and it achieved 96.03% accuracy rate. However, the authors did not mention the size of their test dataset which can provide a strong indication on how generalized their model is.

The work in [9] criticized ROI-based techniques due to their localization limitations and opted to use region-based convolutional neural network technique. The main contributions of [9] are outlined as follows: (1) the annotation of a training images dataset, (2) the training and testing of state-of-the-art object detection techniques which include region-based CNN and exemplar-SVM, and (3) elimination of frequently used assumptions and consideration of practically sound

conditions to design a general solution for LP detection. The proposed methodology first converts the input image into grayscale image, then it uses Gaussian blur filter to reduce noise. The filtered image is then binarized by applying thresholding (Global, adaptive and cluster thresholding techniques), then, the binarized image is searched for contours. Finally, contours are scrutinized for false positives to detect the actual LP. The outcomes of these operations are then fed to machine learning classifier. The accuracy achieved by the technique proposed in [9] is 99.01%.

The authors in [13] proposed license plate detection and recognition using a Histogram of Oriented Gradients (HOG) feature based SVM classifier for clearly visible environments. The training images were downloaded from Google images. Then, the images are fed to HOG SVM detector, and finally, the license plates are extracted. We note that the work in [13] is not suitable for challenging context which might not be helpful in real time environments. Moreover, the authors claim that their classifier achieved good accuracy, however, there is no qualitative or quantitative illustration of results to support their claim.

Saudi Arabia license plate detection based on Artificial Neural Networks (ANN) and objects analysis was proposed in [14]. In the image pre-processing step, several operations are performed: background removal, image cropping and conversion of RGB colored images into grayscale images, and histogram equalization. In the second step, both vertical and horizontal edges are detected and enhanced. In the third step, the vertical and horizontal edges are collected to construct candidate LP objects. Finally, the ANN classifier categorizes the recognized objects as either LP or non-LP. The proposed algorithm achieved a high detection accuracy rate of 98.68%.

A combination of sliding window, HOG, and SVM methods were used to detect license plates in [15]. The developed system is an android-based vehicle plate detection system. Sliding window method is used to determine the location of the plate on the vehicle's image. The HOG method is used for feature extension by taking the texture of the plate. While the SVM method classifies the image to determine whether the detected object is a license plate or not, thus, the location of the plate can be determined. This system achieved accuracy of 96%.

In general, a key challenge for ML-based methods is the process of extracting features before training any model. Feature selection may significantly affect the performance of the ML model [16].

## II. Deep Learning Based Techniques

In this section, we present some of the techniques that employed deep learning for license plate detection. A Convolutional Neural Network (CNN)-based detection scheme is proposed in [17] for complex scenes. The work considers a pipeline of 3 steps. Initially, a simple CNN with 5 convolutional layers is trained. The trained CNN captures the vehicle proposal with the highest Intersection-over-Union (IoU) overlap with a ground-truth box. A second CNN is then used to detect candidate LP from vehicle proposals with high confidence obtained from the first CNN. Finally, the scheme refines the bounding boxes in accordance with edge feature of license plates. For each box side, it enlarges high confident bounding boxes' edges by 30%. The Canny operator is performed on the enlarged region of license plate for edge detection. To evaluate the performance of the proposed scheme, the authors in [17] used dataset of different vehicle types under different real traffic environments. The scheme achieved 90.51% recall rate. We note that the scheme uses 3 step solution, which may require more processing time and slow the detection process.

Another 3 steps scheme using CNN was proposed in [18]. In the first step, a CNN is applied to the vehicles' images to produce convolutional feature maps. The generated maps are then searched for a sub-window with the complete license plate. The search is done using single-scale sliding-window detector. In the third step, a regression network is used to detect the license plate. This scheme achieved 98.4% recall rate.

Another CNN-based scheme for license plate detection is proposed in [19]. It consists of 3 phases. The first phase is image partitioning: it takes the RGB images as an input, then, these images are partitioned into 120 x 180 sub-regions. One LP is only allowed to appear in each sub-region, and the entire LP should fit in the sub-region. The second phase is region processing: each image sub-region is fed to a CNN, in which it is assigned a value in the [0:1] range to indicate whether this sub-region contains a LP or not. The final phase is result integration, where the CNN results are analysed to locate the regions of interest. The proposed system achieved a recall rate of 83%.

The work in [20] proposed a two-step CNN-based solution. The authors in [20] claim that relying on SVM-based solutions only results in many false positives. Hence, their solution uses both deep learning and machine learning techniques. First, they used a deep learning scheme, namely You Only Look Once (YOLO-V2), to detect the whole vehicle in the target image. Secondly, a well-

trained SVM model is used to detect license plates from the detected vehicle from the first step. This solution achieved 94.23% recall rate.

The authors in [21] proposed a deep learning-based scheme for license plate detection in complex scenes. The proposed scheme consists of a two-stage pipeline. The first stage employs a CNN model that consists of 3 convolutional layers, 3 pooling layers and a fully-connected layer. The second stage employs an SVM classifier to detect license plates from the candidate regions obtained in the first stage. This solution achieved 98.81% recall rate.

Vehicle license plate detection and recognition using deep neural networks and generative adversarial networks is proposed in [22]. The detection is based on cascaded CNN. The candidate windows are scanned by the early levels of the CNN to reject the regions that are not considered as license plates. On the other hand, the late layers of the network perform a systematic evaluation of a limited number of candidate windows in high-resolution. Once a plate is detected, it is then cropped and fed to CNN and recurrent neural network for recognition. A connectionist temporal classification layer is used to decode the output results to a readable character sequence. The proposed recognition model is segment-free. The proposed model achieved a high recall rate of 97.95%.

The authors in [23] proposed a license plate detection model in the wild. In [15], YOLO is customized to detect license plates in complex scenes. Two versions of YOLO (YOLO and YOLO9000) are utilized. Both versions customized fully connected layers to split the input images to 11x11 grids to enhance the detection of license plates. Each model was trained with different training set sizes: 1500 instances, 2861, 3161 and 1661. However, all models were tested using the same dataset. The obtained results show that YOLO9000 outperformed YOLO in all cases. The first model achieved unsatisfactory results as the dataset was not representative enough. However, the fourth model achieved the best precision results although the training set is smaller than the one in models 2 and 3 which indicates that the used dataset is adequate for LPD in all aspects.

In [24], a novel three stage LPD system is proposed for detecting English and Chinese license plates. In the first stage, the authors employ an Extremal Regions (ERs) as a character proposal module to detect candidate characters. In the second stage, CNN is used for feature extraction and classification. Finally, region linking is used for license plate detection. Skewed license plates were

detected by the proposed LPD model. When tested with dataset of Chinese vehicles images, the proposed model achieved an accuracy of 98.3%. However, when tested with a dataset of images from 12 different countries, the detection rate was 96%.

The work in [25] proposed a robust license plate recognition for real-time applications based on the YOLO detector. In the detection phase, license plate is detected in 3 steps: in the first phase, a YOLO network is used to detect vehicle. Then, another YOLO network is used to detect LP in the vehicle with the confidence level set to zero. Finally, it picks the object with the maximum confidence as a license plate. For the recognition phase, a CNN is used to segment the characters in the detected LP, obtained from the detection phase. In fact, two CNN networks were trained for the recognition phase. Two different models were trained with SSIG dataset and UFPR-ALPR dataset. The model trained with SSIG dataset achieved a recall ratio of 100%, for both vehicle detection and license plate detection. On the other hand, the model trained with UFPR-ALPR dataset achieved a recall rate of 100% for vehicle detection and 98.33% for license plate detection.

In fact, YOLO was successfully used for object detection. For example, the work in [26] proposed pedestrian detection in video surveillance using fully convolutional YOLO neural network. The network was trained to work with images with arbitrary sizes by using convolutions layers instead of fully-connected layers and train these layers from scratch. However, [26] does not provide any quantitative results to evaluate its performance and accuracy. Another example is introduced in [27], a novel YOLO-based real-time people counting approach is proposed, namely YOLO based People Counting (YOLO-PC). YOLO-PC uses 9\*9 grid and 3 bounding boxes. The authors conducted three sets of experiments using 4 minutes videos obtained from different threshold values (such as 0.2, 0.3 and 0.4). The results show that YOLO-PC is more efficient than the original YOLO in terms of detecting and counting people.

The authors in [28] proposed a method for vehicle number plate recognition. This work reads the colored images and manually crop the license plate region from it. Then, it performs the following operations: analyze the cropped image in its red, green, and blue frames, convert the RGB image to HSV and analyze its Hue, Saturation and Value, apply histogram equalization for image enhancement, use thresholding for image quantization, apply noise reduction, and finally, use segmentation techniques to extract characters and numbers from the image. This work achieved

90% accuracy, however, it is not practical for real time applications as it requires manual cropping of the license plates in the images.

The work in [29] proposed license plate recognition system for vehicle identification in Iran. The paper assumes that the LP is extracted. Then, the input images are converted to greyscale. It performs contrast and brightness adjustment as well as edge detection and median filter on the greyscale images. Once these operations are applied, the LP is ready for segmentation. Characters are segmented, and additional lines are deleted. The segmented characters are then recognized and validated against jurisdiction's specific rules. The paper claims that the proposed system can robustly detect and recognize vehicles using license plate in different lighting and weather conditions. However, no evaluation criteria and results were reported in the paper.

The work in [30] proposed a license plate detection and recognition model. In this model, images are initially converted to greyscale images then converted to binary images using adaptive thresholding. The edge detection is applied next, then, LPs are extracted. The authors did not report any quantitative results to verify the effectiveness of their proposed scheme.

### **III. Font Evaluation**

For font evaluation, we present some of the works in the literature that look at different aspects to evaluate fonts and examine their impact on text readability.

The work in [31] evaluates fonts for digital publishing and display. A vast number of different fonts can be found in various visual styles, these fonts are used in digital publishing and information display. Hence, it is crucial to evaluate their impact on our daily usages. [31] advocates that fonts may have some role in the identification of letters that triggers human word recognition. It is important to view font characteristics as a prospect to enable letter recognition. For example, [31] highlighted the importance of spacing on text reading and word recognition. Spacing typically are used to signal pauses. However, it can also be used to signal grouping and to illuminate the black box of the reading process. The authors proposed a new shape descriptor for ancient characters recognition based on their analysis of the graphic features of inscriptions. The descriptor is translation, rotation, and scaling invariant. The effectiveness of the descriptor to classify characters was confirmed by experimental results. Furthermore, the work in [31] showed that the relationship between font legibility and design features can be better understood using the typography domain.

The authors in [32] investigate the effects of font type and text spacing for online readability and performance. For on screen text, it measures the effects of font type (e.g., serif and sans serif) and spacing as well as its readability to enhance text reading and understanding. The relationship between fonts and readability is established by many researchers. The common serif and sans serif typefaces are best described by Verdana, and Times New Roman fonts, hence, the work in [32] considered both Verdana, and Times New Roman fonts. In the experiments, four 200 word paragraphs with the same level of difficulty and with different font types and different spacing were read by the students. Reading times were measured for each student to identify which font type and spacing is easier to read. The obtained results concluded significant differences in terms of text readability and reading performances between serif (Times New Roman) and sans serif (Verdana) fonts. The results show that sans serifs typeface have better computer screen readability than serif typeface.

The work in [33] proposed a collaboration between type design and science through a new typeface called Sitka to examine font legibility from typographer's point of view. The authors state that in order to enhance the words readability, it is essential to enhance the individual letters recognizability. Hence, they established a new general-purpose typeface. The proposed typeface is designed for on screen usage. The typeface is serif with Roman, Italic and Bold. It supports Latin language as well as Greek and Cyrillic characters. The main goal of the new design is to speed up the evaluation studies and design iterations. Evaluation studies examined different letter forms at different stages in an iterative manner. At each stage, the authors analyzed the obtained results to improve the next stage. This work focus is legibility, in both designing legible typeface and conducting legibility studies.

The obtained results show that a large x-height comes with a trade-off. At one hand, large size is more suitable for neutral height letters, on the other hand, large size is not suitable for ascending and descending letters. The obtained results confirm some of the previous results in the literature. For example, the narrow letters inherent problems.

The authors in [34] considered designing legible fonts for distance reading. It looked at the distance optimality and the legibility of signage fonts. The authors claim that distance has a substantial effect on the readability of smaller features and details. The obtained results show that serifs on the vertical stems enhance distance legibility. However, the low contrast sans serif fonts were

shown to achieve better results in certain cases: the type is influenced by halation, or the light is projected from the back of the letters.

In general, distance legibility tends to improve with open inner counters. The authors proposed to largen the x-height and widen the letter shapes to achieve this goal. Moreover, distance reading was improved by large inter letter spacing.

In [35], the authors examined the usage of Personal Digital Assistants (PDAs) in some Arabic Communities and the suitable fonts used on those devices. They examined 13 typefaces with six different factors. All fonts were set to a fixed size. A normalization process was first done to fonts with different heights while preserving the height and width ratio. Many experiments were conducted to test both letters and words legibility. The experimental results for word legibility were integrated with the letters legibility to form a single component. The results show that the Almohanad font achieved the highest legibility score of 79%. Both Script Hafs and Geeza Pro fonts are recommended for e-books which require high legibility. Due to their good performance with smaller size characters, Almohanad, Geeza Pro, and Yakout Reg fonts are recommended for iPad mini devices with small sized screens.

The authors in [36] examined the effects of both font type and size on fonts legibility and reading time of online text by older adults [36]. They examined both serif and sans serif fonts at 2 different font sizes: 12 and 14 points. In this study, the participants were asked to read online text that contain 10 substitution words (inappropriate grammatically word) which are randomly placed in the text. The legibility is then calculated by the participants accuracy of finding substitution words and the time needed to complete the task. The obtained results show that the 14-point fonts are more legible and promote faster reading. Moreover, the serif 14-point font seems to promote faster reading. In general, sans serif fonts were more preferred than the serif fonts. Furthermore, the results show that there is no difference between the print fonts and the computer fonts.

The authors in [37] examined the legibility of two new ClearType fonts (Cambria and Constantia). The legibility of these 2 fonts was examined against traditional serif font (Times New Roman). The short-exposure method was used in this study. Participants were exposed to a set of letters, digits, and symbols for 34 milliseconds with blanking time of 1.5 seconds. The newly designed fonts reported positive results. The choice of digits used in the Constantia font deemed inappropriate, considering the digits 1, and 2 were confused with the with letters l, and z. However,

the traditional Times New Roman font achieved poor accuracy in particular in digits and symbols results.

## Chapter 3. Proposed Model

License Plate Detection (LPD) in complex scenes is a challenging problem due to its sensitivity to environmental factors (such as rain, dust, and shadow) and lighting conditions, which may greatly influence the detection accuracy. Moreover, LPD is more challenging for real time systems.

Furthermore, the chosen font type in the license plate plays a vital role in the recognition phase in computer-based studies. Despite its vital importance and its great effect on the recognition accuracy, font evaluation for recognition is not sufficiently investigated in the literature.

Hence, this work is two folded. On one hand, we propose to employ Deep Learning (DL) techniques in license plate detection. On the other hand, we propose to evaluate font characteristics in the license plate context.

The main contributions of this work are:

1. Examine the effect of customizing important YOLO hyperparameters to improve the detection recall rate for license plate detections.
2. Perform a qualitative analysis of the false predictions of the trained YOLO models.
3. Analytically evaluate 2 fonts used in license plates and investigate how the typeface design can affect computer-based studies in context.

The rest of this chapter is organized as follows: the datasets used in this study are introduced in Section I. Section II formulates the license plate detection problem and the use of YOLO framework. The proposed detection model is presented in Section III, and the font evaluation study is presented in Section IV.

### I. Datasets

Deep learning techniques depend heavily on the available dataset and its size. To conduct the experiments of this research, two datasets were used:

1- Federal University of Paraná dataset for Automatic License Plate Recognition (UFPR-ALPR) [25]. The dataset consists of 4500 images of license plates from Paraná /Brazil. This dataset was captured with three different cameras; GoPro Hero4 Silver, Huawei P9 Lite and iPhone 7 Plus. It

consists of images taken from motorcycles, cars and public transportation vehicles. The main characteristics of this dataset are:

- a) The dataset is split by UFPR as follows: 40% for training, 20% for validation and 40% for testing.
- b) All images in this dataset are taken for moving vehicles during daytime.
- c) The distance between the camera and the license plate varies from 1 meter to 10 meters.
- d) Thirty percent of the images are exposed to bright sunlight, shade, and dust.

Figure 4 presents sample images from this dataset. Figure 5 shows the distribution of letters and digits in UFPR-ALPR dataset. The counts for all letters and digits are shown in Table 2.



Figure 4. Sample Images from UFPR-ALPR Dataset

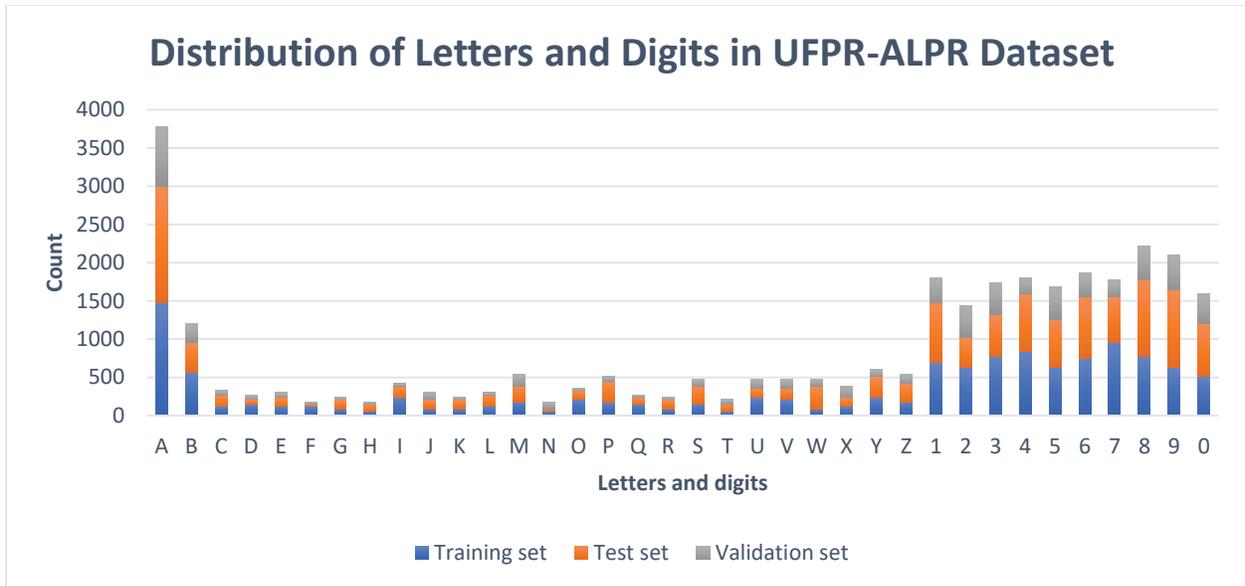


Figure 5. Letter and Digit Distributions in UFPR-ALPR

2- Centre for Pattern Recognition and Machine Intelligence (CENPARMI) dataset of license plates in complex scenes. This dataset consists of 480 images of license plates from Quebec/Canada. This dataset was captured with an iPhone 6 camera. The images for this dataset were collected and annotated by the authors of this thesis. It consists of images taken from motorcycles, cars and public transportation vehicles. The diversity of image brightness setting, vehicle distance, license plate sizes and license plate background color are challenging factors for license plates detection. The main characteristics of this dataset are:

- a) Approximately, 50% of the images in this dataset are taken for moving vehicles and the rest of the images are taken for stationary vehicles.
- b) Eighty percent of the images are taken at daytime, while 20% are taken at parking lots with little or no light.
- c) The distance between the camera and the license plate varies from 1 meter to 30 meters.
- d) In 20% of the images, LPs are pictured at an angle which makes its detection and recognition more challenging.
- e) Twenty percent of the images are exposed to bright sunlight, and shade.

Figure 6 presents sample images from this dataset. Figure 7 shows the distribution of letters and digits in the CENPARMI dataset.

Table 2. Distribution of Letters and Digits in UFPR-ALPR Dataset

Letters/Digits	Training set	Test set	Validation set	Total
A	1470	1530	780	3780
B	570	390	240	1200
C	120	150	60	330
D	150	60	60	270
E	120	120	60	300
F	120	30	30	180
G	90	120	30	240
H	60	90	30	180
I	240	150	30	420
J	90	120	90	300
K	90	120	30	240
L	120	150	30	300
M	180	210	150	540
N	60	30	90	180
O	210	120	30	360
P	180	270	60	510
Q	150	90	30	270
R	90	120	30	240
S	150	240	90	480
T	60	90	60	210
U	240	120	120	480
V	210	150	120	480
W	90	300	90	480
X	120	120	150	390
Y	240	270	90	600
Z	180	240	120	540
1	690	780	330	1800
2	630	390	420	1440
3	780	540	420	1740
4	840	750	210	1800
5	630	630	420	1680
6	750	810	300	1860
7	960	600	210	1770
8	780	990	450	2220
9	630	1020	450	2100
0	510	690	390	1590

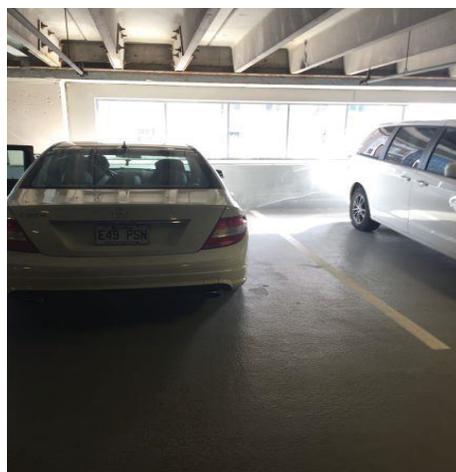


Figure 6. Sample Images from CENPARMI Dataset

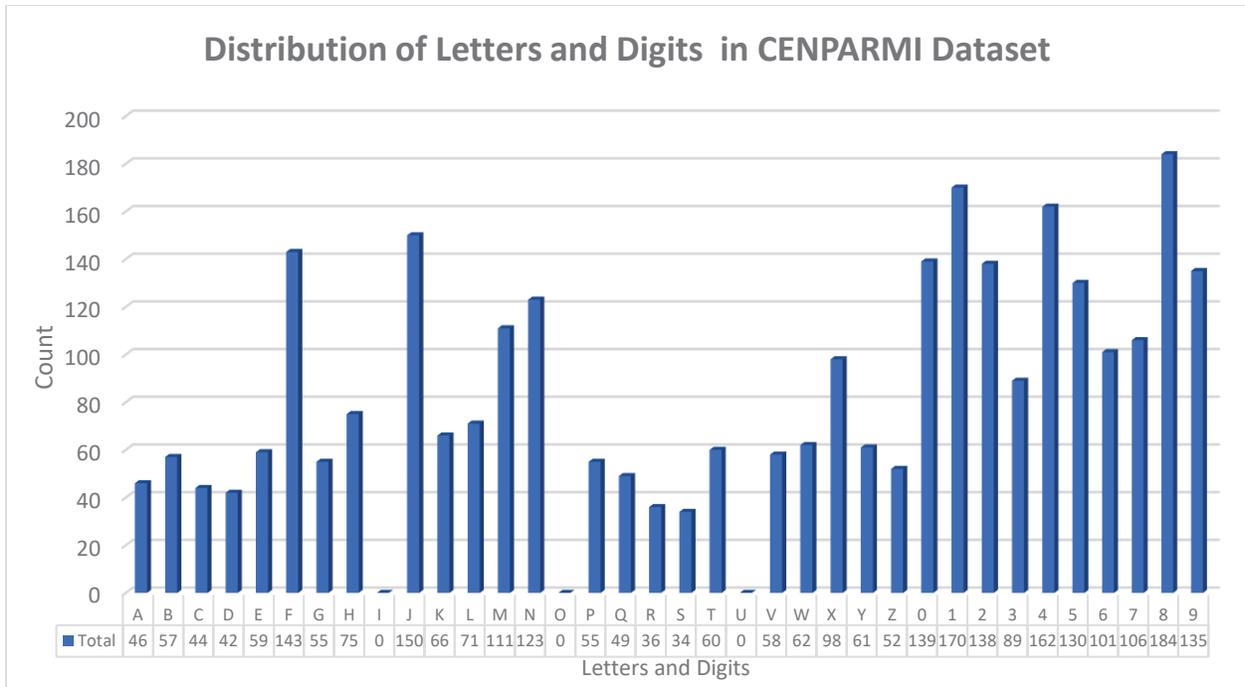


Figure 7. Letter and Digit Distributions in CENPARMI Dataset

## II. License Plate Detection

The usage of deep learning attracted the attention of researchers in recent years. It is being widely employed to solve classification and detection problems. It is used in many domains such as object detection, natural language processing, and computer aided diagnosis. For object detection, many techniques have been developed such as Convolutional Neural Network (CNN). However, in the past 3 years, a new model was developed, namely You Only Look Once (YOLO) [1]. YOLO is a real time object detection system that performs objects' localisation and objects' classification in one scan rather than performing these 2 operations in 2 or more scans.

## III. YOLO

YOLO is a single convolutional network which simultaneously finds multiple bounding boxes and class probabilities for those boxes. YOLO [38] scans the whole image as input, each input image is then divided into an  $S \times S$  grid. Each grid cell predicts  $B$  bounding boxes, where each bounding box consists of 5 predictions:  $(x, y, h, w, confidence)$  where,

$x$ , and  $y$  are the coordinates of the center of the bounding box,

$h$ , and  $w$  are the height and the width of the bounding box relative to the whole image,

and *confidence* is the Intersection over Union (IoU) between the predicted box and any ground truth box. Each grid cell also predicts  $C$  conditional class probabilities.

YOLO predicts multiple bounding boxes per grid cell. At training time, YOLO aims to have one bounding box predictor to be responsible for each object. It assigns one predictor “responsible” for predicting an object based on which prediction has the highest current IoU with the ground truth. This leads to specialization between the bounding box predictors.

This work uses YOLOv2 [1] which is an enhanced version of the original YOLO. YOLOv2 borrows the anchors concept from Faster R-CNN to represent parameterized proposals. A scale and aspect ratio are assigned for each anchor along with hand-picked priors that are attained from the training set. To predict bounding boxes and their center locations, YOLOv2 exploits these anchor boxes.

#### **IV. Proposed Detection Methodology**

The success of deep learning models depends heavily on the available training set. This work uses (UFPR-ALPR) to train all models. later UFPR-ALPR test set and CENPARMI license plates in complex scenes dataset are used for testing.

Since YOLO is a new technique, the choice of certain hyperparameters can affect the detection accuracy. Hence, we propose to examine YOLO hyperparameters and then customize it to enhance its performance in detecting license plates. The following set of experiments will examine different YOLO models with different hyperparameters to study its effect on the predictions results. Moreover, we propose to train the YOLOv2 net using pre-trained weights on Common Object in Context (COCO) dataset instead of randomly initialized weights [2]. Throughout training, a batch size of 64, a momentum of 0.9 and learning rate of 0.001 are used.

All experiments are made on the supercomputer Helios from Université Laval, managed by Calcul Québec and Compute Canada [39]. A total of 32 experiments will be conducted to examine the effect of the following aspects:

1- Input network resolution: this set of experiments will examine the effect of training YOLO net with different input network resolutions for both training and testing nets separately. We propose to examine the following resolution values: 320 x 320, 608 X 608, 800 X 800, and 960 X 960.

Each training input resolution will be evaluated against all 4 resolutions in the testing phase. This study considers low, medium and high resolutions value. The standard input network resolutions used by YOLOv2 is 608, and the maximum possible network resolution on the available resources is 960. The 800 resolution is examined as an average resolution of 608 and 960 (the resolution size must be divisible by 32). The 800 and 960 represent high input network resolution, the 608 represents a medium input network resolution. To examine low resolution scenarios, an input network resolution of size 320 is chosen.

2- Size of training set. This set of experiments will examine the effect of the training dataset size on the obtained predictions. We propose to use the following dataset sizes: 1800 instances from UFPR-ALPR training set, and 2700 (training + validation set of UFPR-ALPR). Each size will be examined for all input network resolution values from step 1. Figure 8 illustrates the steps of the proposed LP YOLO-based detection algorithm.

<b>Algorithm:</b> Proposed LPD
<b>Input:</b> Training dataset (i.e. raw images of cars)
<b>Step 1:</b> Feed the dataset to a customized YOLOv2 network to detect all possible license plates with confidence threshold set to 13%.
<b>Output:</b> Set of detected license plates

Figure 8. Proposed LP Detection Algorithm

The proposed network consists of 22 convolution layers and 5 maxpooling layers. Convolution layers are responsible for extracting features using kernels of sizes 3 x 3 and 1 x 1. At each convolution layer, the filter maps slide over the pixel matrices by 1 pixel (stride) to extract feature maps. In the maxpooling layers, the dimensionality is reduced by choosing the maximum value from every 2 x 2 window in the feature maps.

To illustrate the main structure of the designed YOLO networks, we now present two examples of our designed YOLO networks. Table 3 demonstrates the structure of a YOLO network that is designed to be trained on the network input resolution of value 800 x 800, and Table 4 displays a YOLO network designed for the testing phase, with the value of the input network resolution set to 960 x 960. For the remaining 30 experiments, YOLO networks of the same structure were used with different dimensions for the convolution and maxpooling layers. For example, the size of the first convolution layer in the YOLO network trained with input network resolution 320 x 320, is 320 x 320 x 3, then it is down sampled to 160 x 160 x 32 in the maxpooling layer and so on.

This work presents a novel methodology of enhanced YOLO networks, to our knowledge, this methodology is not used in the literature. Moreover, this work conducts an analytical study on the effect of tuning the input network resolution. In addition, all models are tested on a confidence threshold of 13% that is obtained empirically by running multiple experiments on the validation set of UFPR-ALPR.

Table 3. Training YOLO network structure

	Layer	Filters	Size	Output size
1	Input			800 x 800 x 3
2	Conv	32	3x3/1	800 x 800 x 32
3	maxpooling		2x2/2	400 x 400 x 32
4	Conv	64	3x3/1	400 x 400 x 64
5	maxpooling		2x2/2	200 x 200 x 64
6	Conv	128	3x3/1	200 x 200 x 128
7	Conv	64	1x1/1	200 x 200 x 64
8	Conv	128	3x3/1	200 x 200 x 128
9	maxpooling		2x2/2	100 x 100 x 128
10	conv	256	3x3/1	100 x 100 x 256
11	conv	128	1x1/1	100 x 100 x 128
12	conv	256	3x3/1	100 x 100 x 256
13	maxpooling		2x2/2	50 x 50 x 256
14	conv	512	3x3/1	50 x 50 x 512
15	conv	256	1x1/1	50 x 50 x 256
16	conv	512	3x3/1	50 x 50 x 512
17	conv	256	1x1/1	50 x 50 x 256
18	conv	512	3x3/1	50 x 50 x 512
19	maxpooling		2x2/2	25 x 25 x 512
20	conv	1024	3x3/1	25 x 25 x 1024
21	conv	512	1x1/1	25 x 25 x 512
22	conv	1024	3x3/1	25 x 25 x 1024
23	conv	512	1x1/1	25 x 25 x 512
24	conv	1024	3x3/1	25 x 25 x 1024
25	conv	1024	3x3/1	25 x 25 x 1024
26	conv	1024	3x3/1	25 x 25 x 1024
27	concat		[16]	50 x 50 x 512
28	conv	64	1x1/1	50 x 50 x 64
29	local flatten		2x2	25 x 25 x 256
30	concat		[27, 4]	25 x 25 x 1280
31	conv	1024	3x3/1	25 x 25 x 1024
32	conv	30	1x1/1	25 x 25 x 30



IoU is the result of dividing the intersection between the predicted Bounding Box (BB) and the Ground Truth (GT). For false positives, we classify IOU values into three classes:

- A.  $\text{IoU} = 0$ : Since the union of BB and GT is always a positive number, IoU will be zero if and only if the intersection of them is zero. This might happen in the following cases:
  - a. The GT is not correctly annotated in the test set. For example, UFPR-ALPR has more than thirty unannotated license plates in its test set.
  - b. The model predicted an irrelevant object.
- B.  $0 < \text{IoU} < \text{threshold}$ : indicates that the object is correctly detected with shifted or insufficient localisation.
- C.  $\text{IoU} \geq \text{threshold}$ : redundant detection (i.e., the same true positive object gets detected more than once), hence, it is considered as false positives.

## V. Digital Font Evaluation

Automatic License Plate Recognition (ALPR) consists of 3 main steps: detection, segmentation, and recognition. We note that even the best segmentation algorithms do not always produce recognizable letters and digits from the cropped segments, hence, each character must maintain unique integrity to overcome cropping shortcomings. Therefore, choosing the adequate font is a key to successful license plate recognition, even for general plate numbers that do not support regular expressions. This work proposes the following contributions:

1. Study the characteristics of 2 different fonts: Mandatory, and the font used for license plates in Quebec. Fourteen different font cases will be considered: A and 4, B and 8, C and G, D and P, E and F, I and 1, 8 and 6, 0 and O, Q and O, U and V, V and Y, Z and 2, S and 5, G and 6 designs. Each glyph was taken individually and inspected in context anatomy to obtain both positive and negative characteristics that may affect its recognizability. The anatomy study will consider the aspects listed in Table 5.
2. Evaluate the effect of the used fonts in context for 2 datasets: UFPR-ALPR dataset for the Mandatory font, and Quebec license plate dataset for Quebec license plate font. The evaluation will be conducted using 2 commercial products: OpenALPR [4], and Plate Recognizer [5]. Two sets of results will be attained. The first set will report the recall values while considering the entire predicted license plate as true or false. On the other hand, the second set of results reports more qualitative results; it reports confusion matrices for each

letter in each license plate. Confusion matrix will help us to identify the level recognizability of each character.

Based on the obtained results, we propose to produce a set of recommendations for adequate font characteristics to enhance recognition accuracy in real time.

Table 5. Font Type Evaluation Aspects

Aspect	Explanation
Similar apex	Apex is defined as the juncture of a stem. If two glyphs have the apex at the same position and design, the probability of confusion will increase. For example, the top apex in letter A and digit 4.
Similar crossbar position	Crossbar is defined as the horizontal stroke connecting two parts of a letter form (e.g. H). If two glyphs have crossbars, then locating the crossbars at different positions would improve their recognizability.
Similar top counter	Counter is defined as the negative space fully or partially enclosed by the stroke of a letter form. If two glyphs have top counters that look identical, probability of confusion will increase.
Similar bottom counter	If two glyphs have bottom counters that look identical, probability of confusion will increase.
Similar bowl	Bowl is defined as the fully or modified rounded forms. If two glyphs have the bowl at the same position and design, the probability of confusion will increase.
Identical spur	Spur is defined as a projection smaller than a serif, that reinforces the point at the end of curved stroke. If two glyphs have the spur at the same position and design, the probability of confusion will increase.
Identical bottom, or top horizontal stroke	Stroke is defined as any linear element that makes up a letterform. If two glyphs have the stroke at the same position and design, the probability of confusion will increase.
Similar diagonal stroke	If two glyphs have the diagonal stroke at the same position and design, the probability of confusion will increase (e.g., 5 and S).
Tail not clear	Tail is defined as a short downward stroke. If the tail is not clear or visible in context, it is more likely to be confused with other glyphs.

# Chapter 4. LPD Results and Analysis

Here, we present the results of LPD models and error analysis.

## I. LDP Results

The recall results for YOLO model trained with 2700 and 1800 instances are shown in Figures 9 and 10 respectively. Each line in the figures corresponds to a specific training network input resolution. The examined test network input resolutions are: the blue line (diamond shape) corresponds to 320, the orange line (square shape) corresponds to 608, the grey line (triangular shape) corresponds to 800, and the yellow line (cross shape) corresponds to 960.

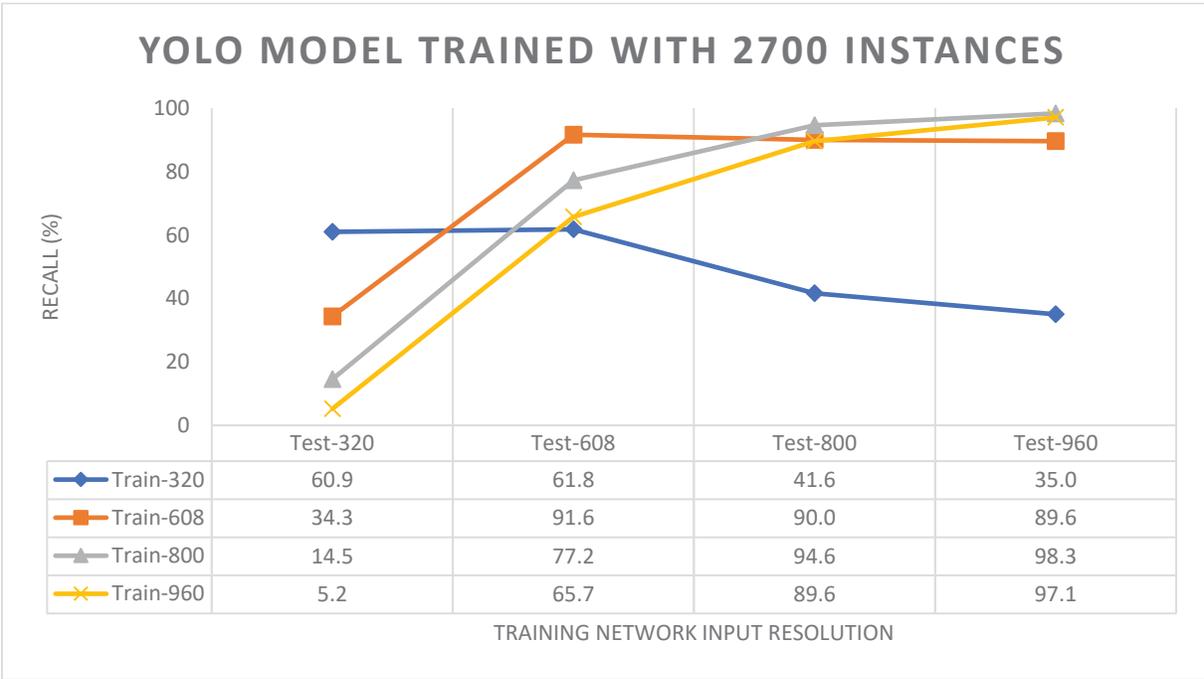


Figure 9. Recall Results for YOLO Models Trained with 2700 Instances

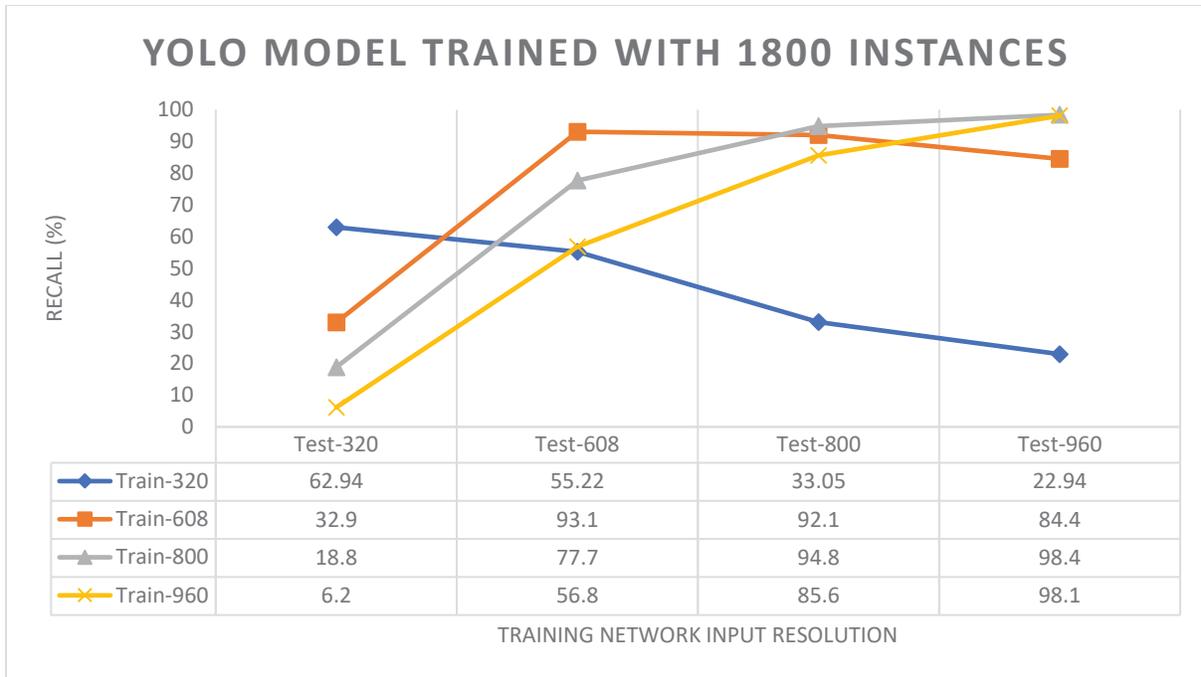


Figure 10. Recall Results for YOLO Models Trained with 1800 Instances

From Figures 9 and 10, we can make the following observations:

1. For low training network input resolution (blue line 320), the recall value decreases as the test network input resolution increases. And the highest recall value is about 60%, which is considered as unsatisfactory recall rate. For these models, the network is trained with low resolution, and increasing the test resolution degraded the model performance as the new high resolutions are not recognized (trained) by the model.
2. Training the network with low input resolution and testing it with a higher input resolution yields very poor recall values.
3. For medium training network input resolution (orange line 608), the best recall value was achieved at the matching training resolution of 608. Increasing the test network input resolution resulted in slight decreases of the recall ratio. At 320 training input network resolution, the recall ratio was very low (about 32%).
4. For both resolutions, 320 and 608, the best recall ratio was achieved when the model was tested with the matching test resolution. At these two training resolutions, the network failed to recognize the license plate features once the license plates are tested at different sizes, this behaviour is considered as drawback for any deep learning model. Moreover, we

note that as the network resolution increases (higher than the training resolution), the recall ratio decreases.

5. Model trained at high network input resolutions (800) achieved the best recall ratios which indicates a well generalised model and well learnt features, and the recall increases as the test network resolution increases. At this network input resolution, the model achieved its best recall at a higher test input resolution (960) which reflects robustness of this model and its ability to detect license plates at a higher resolution than its trained resolution.
6. Model trained at the highest input resolutions (960) achieved high recall ratio at high test resolutions. The recall ratios were slightly lower than the ones achieved at a resolution of 800.
7. From points 5 and 6, we conclude that increasing the training network input resolution does not necessarily warrant an increase in the recall ratios.
8. Both dataset sizes (2700 and 1800) achieved comparable recall ratios, hence increasing the training dataset size beyond certain threshold may not be beneficial for recall ratio.

Since the model trained with resolution of 800 showed robustness and achieved the best recall ratio, it was tested with another dataset (CENPARMI dataset), the images examined in this dataset are taken by different cameras with different resolutions. The model achieved 94.7% recall ratio. This result demonstrates the generality and the robustness of this model.

The performance of our proposed model is compared against the performance of the model proposed in [25]. The model in [25] is trained on UFPR-ALPR dataset, to detect LPs of vehicles and motorcycles, it uses two YOLO networks for detection. The first YOLO network deploys 12 convolutional layers and 3 max pooling layers to detect vehicles and motorcycles. Then the detected objects from the first network are cropped before being fed to the second YOLO network. The second YOLO network deploys 22 convolutional layers and 5 max pooling layers. The second network emits all potential LP objects by setting its confidence threshold to zero. Then, the emitted LPs are fed to a function that selects the license plate with the highest confidence score as a final output. The detection in [25] is done in 4 phases. Whereas, the detection in our proposed model is done in one step only using one YOLO network with 22 convolutional layers and 5 max pooling layers. Hence, the computation power needed for our model is improved and achieved similar

recall ratio of 98.3%. The main conclusion of our detection algorithm is that tuning YOLO's hyperparameters can lead to more efficient design of detection models.

To further elaborate on the obtained results for the proposed model and the model in [25], we make the following additional observations and remarks:

- a) The model in [25] uses two YOLO networks for detection which increases the training and the detection efforts and the computational costs. The design in [25] contradicts with the general intuition of YOLO which emphasizes on eliminating the pipelining process in the detection models. Moreover, the model in [25] needed about 80,000 iterations to train the model, whereas, only 7000 iterations were needed to train our model.
- b) The confidence threshold used in [25] for detection is set to 0% which is a risky value and means that the trained model is 0% confident about the results and might indicate that the network needs more training or extra layers. Whereas, the confidence threshold used in our model is set to 13%.
- c) For the model in [25], the outcomes of the second phase are filtered using a maximum function. Hence, the predictions of the model are not fully automated by the neural networks.

Figures 11 and 12 present sample images from the detection results from CENPARMI and UFPR-ALPR respectively.

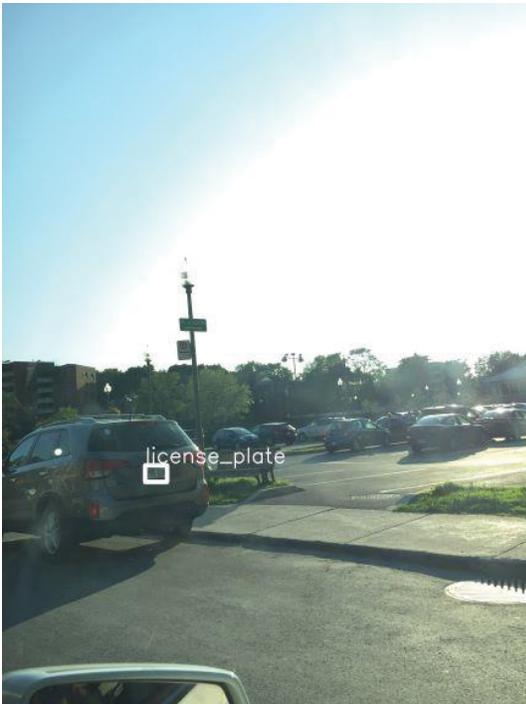
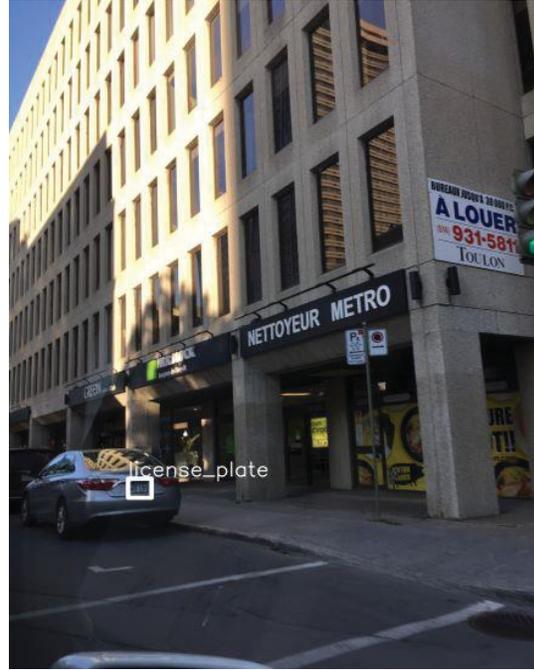


Figure 11. Sample Detection Results for the YOLO Network Trained with 800 Resolution and Tested with 960 Resolution from CENPARMI Dataset



Figure 12. Sample Detection Results for the YOLO Network Trained with 800 Resolution and Tested with 960 Resolution from UFPR-ALPR Dataset

Moreover, we compared our results with those obtained from two commercial systems: OpenALPR and Plate Recognizer. The detection results were obtained for the detection of license plates with UFPR-ALPR test dataset. For the commercial products, we extracted the detection results only, recognition results are not considered as this part is concerned with the detection models only. The detection recall ratios are listed in Table 6. The recall ratio results for the proposed model and the ones obtained from the two commercial products show the superiority of the proposed detection model over the existing commercial products.

Table 6. Comparison of Recall Ratio Results

Model/System	Recall Ratio
Proposed model	98.38%
Open ALPR	64.6%
Plate Recognizer	38.1%

Intuition suggests that with deep learning and increasing the dataset size will improve the model performance. However, as shown in Figures 9 and 10, both datasets achieved comparable recall ratio results. Thus, to further clarify the effect of different data set sizes on the training model, we present and discuss another performance metric: number of iterations attained to train the model. Figure 13 depicts the number of iteration results for both data sizes (2700, and 1800), it shows that models with a bigger dataset size (i.e., 2700) required less number of iterations to be trained (i.e. less training time) than models with smaller dataset size (i.e. 1800). Thus, they required less time to reach comparable recall ratios.

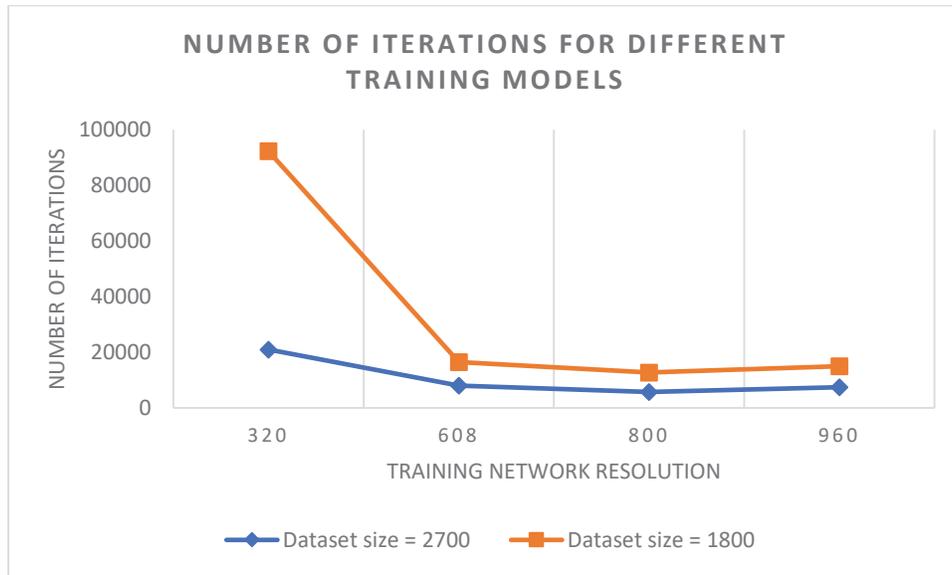


Figure 13. Number of Iterations for Different Training Models

## II. Error Analysis

To further understand the nature of errors occurred during the license plate detection process using YOLO, we identified 3 types of possible errors: wrong object, wrong localization, and redundant detection. The IoU value is used to determine each error type. The error analysis based on error

type is presented in Figures 14 to 21. Each figure shows the percentage of errors resulted from each type.

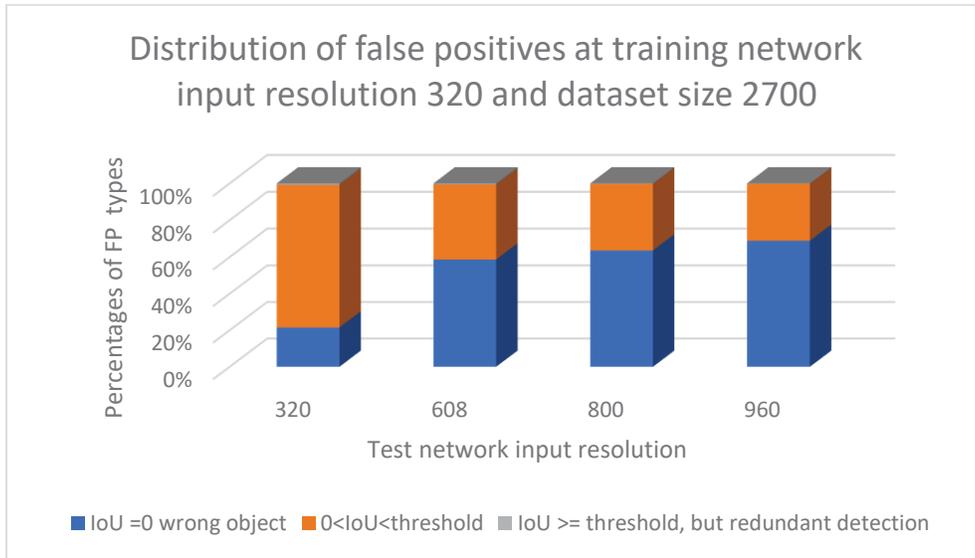


Figure 14. Distribution of False Positives at Training Network Input Resolution 320 and Dataset Size 2700

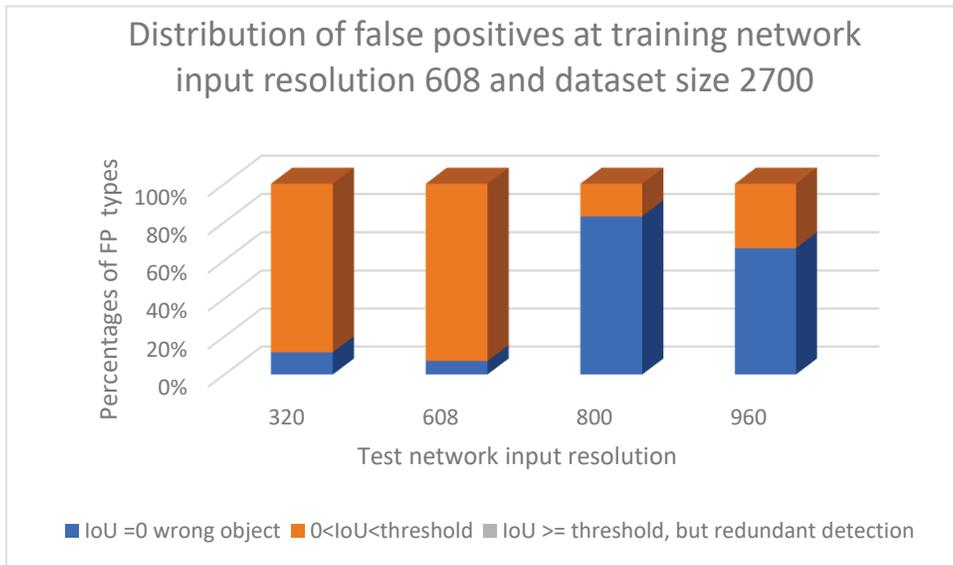


Figure 15. Distribution of False Positives at Training Network Input Resolution 608 and Dataset Size 2700

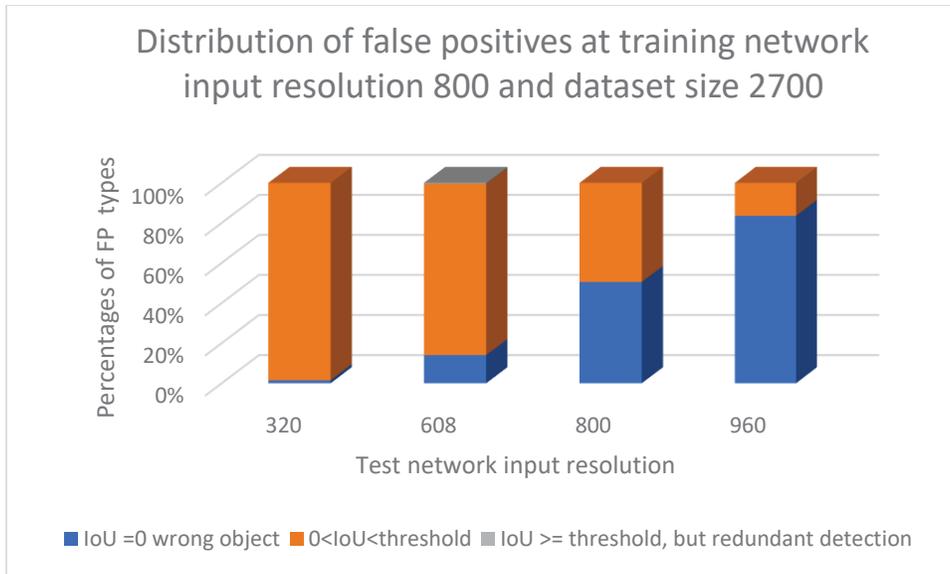


Figure 16. Distribution of False Positives at Training Network Input Resolution 800 and Dataset Size 2700

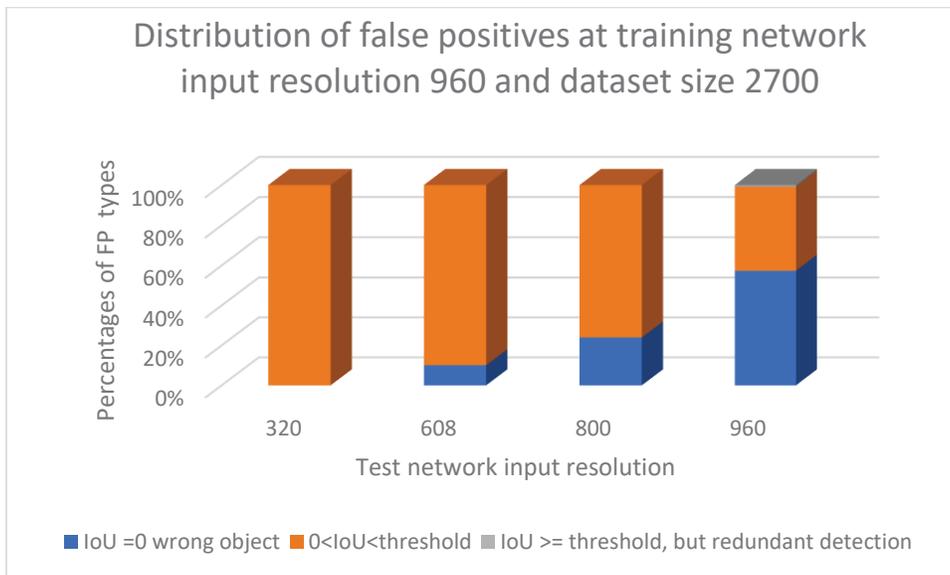


Figure 17. Distribution of False Positives at Training Network Input Resolution 960 and Dataset Size 2700

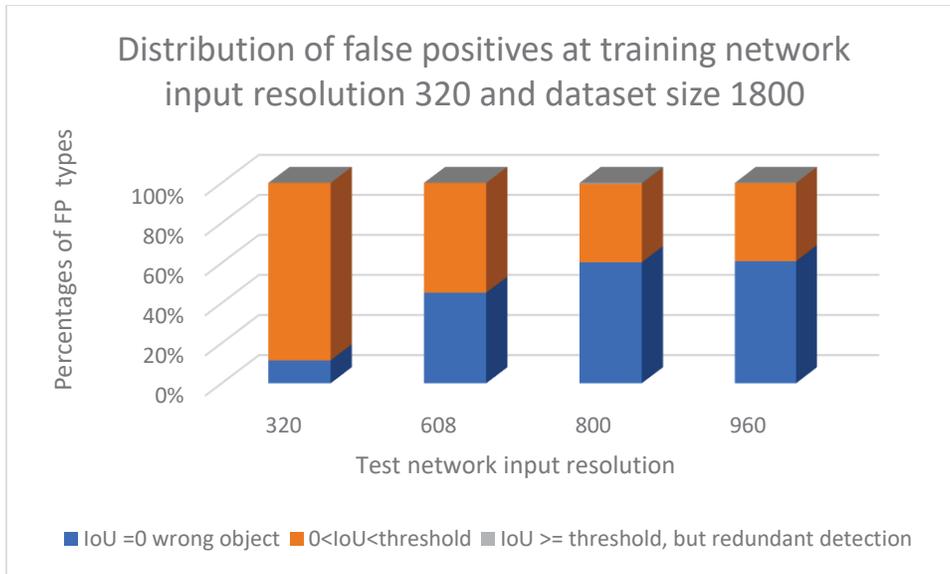


Figure 18. Distribution of false positives at training network input resolution 320 and dataset size 1800

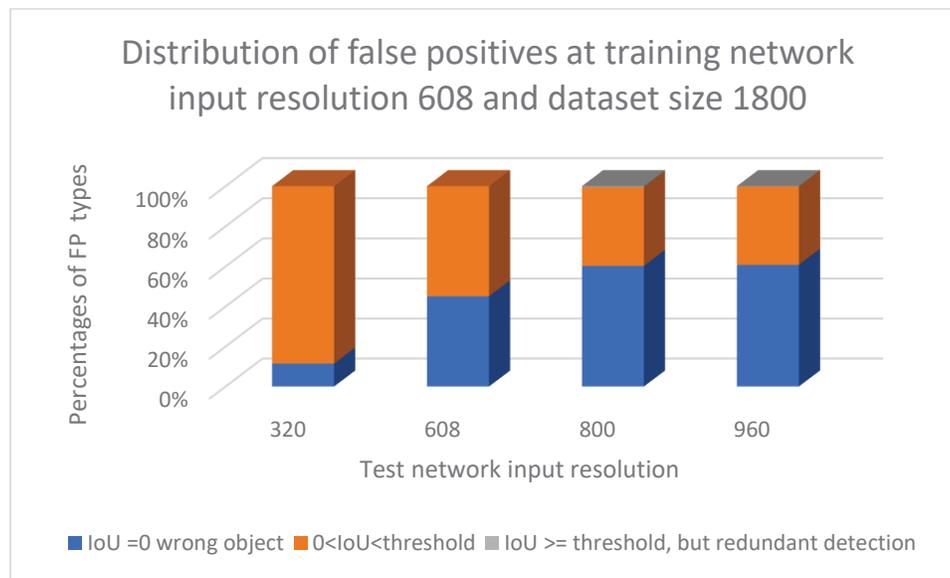


Figure 19. Distribution of False Positives at Training Network Input Resolution 608 and Dataset Size 1800

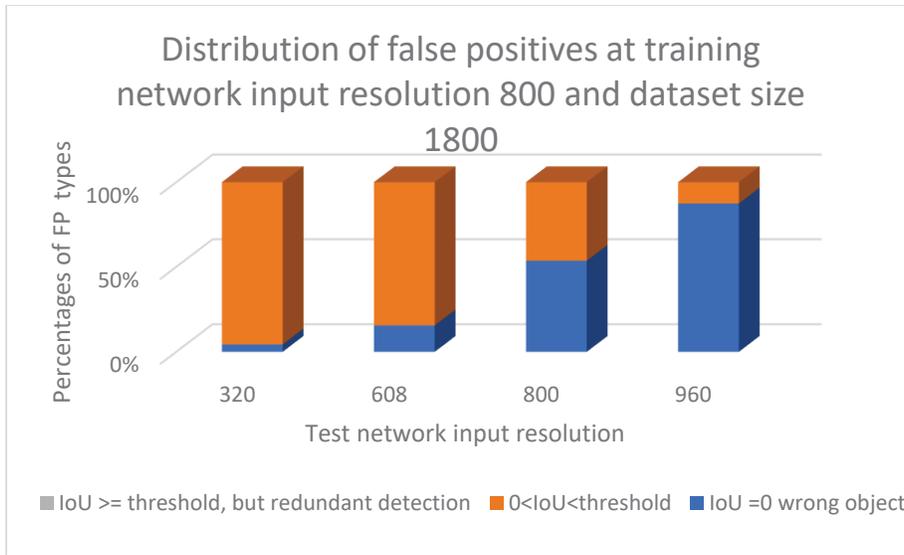


Figure 20. Distribution of False Positives at Training Network Input Resolution 800 and Dataset Size 1800

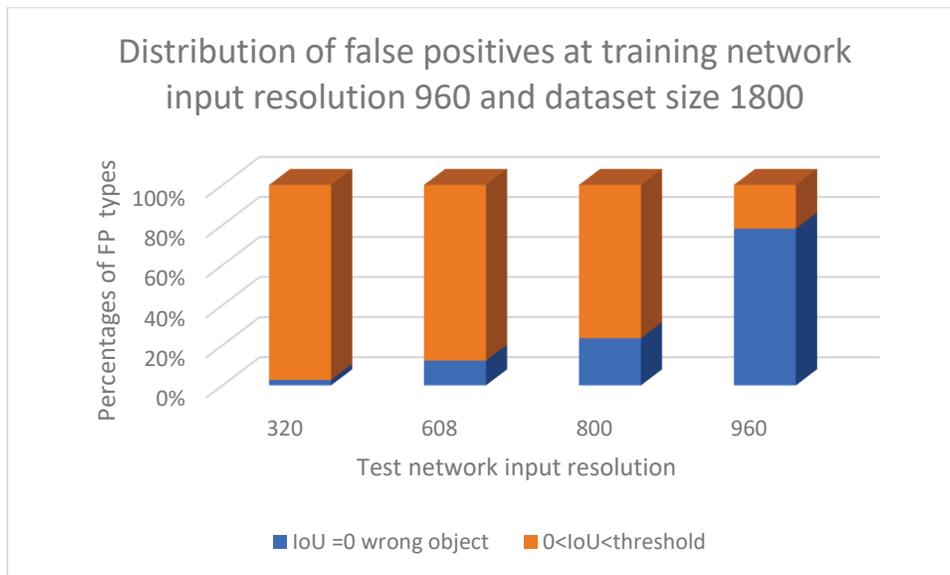


Figure 21. Distribution of False Positives at Training Network Input Resolution 960 and Dataset Size 1800

From Figures 14 to 21 we can make the following observations:

1. For all cases, we notice insignificant redundancy errors, even more, at higher training input resolution no redundancy errors are reported.
2. For all models, as the test resolution increases, the localisation errors significantly decrease.

3. For all models, as the test input resolution increases, the irrelevant objects error increases. However, we note that the overall number of false positives decreases.

Figures 22, 23, and 24 present some sample images from detection errors.



Figure 22. Sample Image for Detecting Wrong Object



Figure 23. Sample Image for a LP Considered as FP Caused by Dataset Lack of Annotation



Figure 24. Sample Image for Error Caused by  $0 < \text{IoU} < \text{Threshold}$

## Chapter 5. Font Evaluation Results and Analysis

Here, we present 2 sets of results for font evaluation: font anatomy results, and commercial products recognition results. For anatomy results, 2 fonts are considered: Mandatory, and Quebec. The mandatory font glyphs are available publicly as depicted in Figure 25 [40].



Figure 25. Mandatory Font Glyphs

To conduct the font anatomy study in context, we first collected all images that have two confusion glyphs (e.g., the images that have both A and 4). Table 7 shows the anatomy results for Mandatory font for all 14 cases. Eleven different features were examined and reported for each case.

Table 7. Mandatory Font Anatomy

	Similar Apex	Similar crossbar position	Similar top counter	Similar bottom counter	Similar bowl	Identical spur	Identical bottom horizontal stroke	Identical top horizontal stroke	Similar diagonal stroke	Identical glyphs	Tail not clear
<b>A vs 4</b>	✓	✓	✓	×	NA	NA	NA	NA	✓	×	NA
<b>B vs 8</b>	NA	NA	✓	✓	✓	NA	✓	✓	NA	×	NA
<b>C vs G</b>	NA	NA	✓	NA	✓	×	✓	✓	NA	×	NA
<b>D vs P</b>	NA	NA	×	NA	✓	NA	NA	✓	NA	×	NA
<b>E vs F</b>	NA	✓	NA	NA	NA	NA	✓	✓	NA	×	NA
<b>I vs 1</b>	NA	NA	NA	NA	NA	NA	NA	NA	NA	✓	NA
<b>8 vs 6</b>	NA	NA	✓	✓	✓	NA	✓	✓	NA	×	NA
<b>0 vs O</b>	NA	NA	✓	NA	✓	NA	✓	✓	NA	✓	NA
<b>Q vs O</b>	NA	NA	✓	NA	✓	NA	✓	✓	NA	×	✓
<b>U vs V</b>	×	NA	×	NA	NA	NA	NA	NA	×	×	NA
<b>V vs Y</b>	✓	NA	✓	NA	NA	NA	NA	NA	✓	×	NA
<b>Z vs 2</b>	NA	NA	NA	NA	×	NA	✓	×	×	×	NA
<b>S vs 5</b>	NA	NA	✓	✓	×	NA	✓	✓	✓	×	NA
<b>G vs 6</b>	NA	NA	✓	✓	✓	✓	✓	✓	NA	×	NA

From the analysis presented in Table 7, we can make the following remarks:

- All cases (except Z-2, D-P and U-V) have the potential to cause confusion at recognition.
- The I-1, Q-O, and 0-O cases have high potential for confusion at recognition due to their high geometrical similarities

For the font of Quebec license plates, the anatomy results are shown in Table 8 for all 14 cases. We note that the glyphs are not publicly available. We note that, in non-personalized plates, the letters O, U and I are not used, hence, the confusion cases: O and 0, O and Q, U and V, and I and 1 are not possible. Moreover, for personalized plates, only letter O is not used, hence, the confusion cases: O and 0, O and Q are not possible.

Table 8. Quebec Font Anatomy

	Similar Apex	Similar crossbar position	Similar top counter	Similar bottom counter	Similar bowl	Identical spur	Identical bottom horizontal stroke	Identical top horizontal stroke	Similar diagonal stroke	Identical glyphs	Tail not clear
A vs 4	✓	✓	✓	×	NA	NA	NA	NA	✓	×	NA
B vs 8	NA	NA	×	×	×	NA	✓	✓	NA	×	NA
C vs G	NA	NA	✓	NA	✓	×	✓	✓	NA	×	NA
D vs P	NA	NA	×	NA	✓	NA	NA	✓	NA	×	NA
E vs F	NA	✓	NA	NA	NA	NA	✓	✓	NA	×	NA
I vs 1	NA	NA	NA	NA	NA	NA	NA	NA	NA	×	NA
8 vs 6	NA	NA	✓	✓	✓	NA	✓	✓	NA	×	NA
0 vs O	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Q vs O	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U vs V	×	NA	×	NA	NA	NA	NA	NA	×	×	NA
V vs Y	✓	NA	✓	NA	NA	NA	NA	NA	✓	×	NA
Z vs 2	NA	NA	NA	NA	×	NA	✓	×	×	×	NA
S vs 5	NA	NA	✓	✓	×	NA	✓	×	×	×	NA
G vs 6	NA	NA	✓	✓	✓	✓	✓	✓	NA	×	NA

From the analysis presented in Table 8, we can make the following remarks:

- All cases (except the not applicable ones) have the potential to cause confusion except B-8, Z-2, I and 1, D-P, U-V.
- Number 1 has a top arm that helps to distinguish it from letter I (this does not show in Table 8).

We note that not all anatomy features have the same influence on these cases. One feature (or a combination of some particular features) might increase the confusion between 2 glyphs more than other features.

Now, we present the recognition results for both OpenALPR and Plate recognizer commercial products. We show the recall ratio results as well as the confusion matrices for all characters in all plates. Table 9 presents the recall ratio results for both commercial products using both datasets.

Table 9. Recall Ratio Results

Product name and Dataset	Recall ratio
OpenALPR with UFPR-ALPR	55.88 %
OpenALPR with CENPARMI	51.87 %
Plate recognizer with UFPR-ALPR	23.69 %
Plate recognizer with CENPARMI	50.20 %

Both products achieved relatively low recall ratios. These results are due to the products' shortcomings in detecting and recognizing license plates of motorcycles and license plates with non-white background colors. We note that the OpenALPR achieved a better performance than the Plate Recognizer. And it achieved a higher ratio using the UFPR-ALPR dataset.

Figures 25 to 32 depict the main results of the confusion matrices for both products and datasets. These figures present more details on the recognition behaviour and shed the light on the problematic confusion cases. Each figure plots the stacked percentages of both correct and misclassifications for each letter and digit. Figures 26 and 27 depict the results for plate recognizer with UFPR-ALPR dataset and its confusion matrix heat map, respectively. Figures 28 and 29 depict the results for OpenALPR with UFPR-ALPR dataset and its confusion matrix heat map, respectively. Figure 30 and 31 depicts the results for OpenALPR with CENPARMI dataset and its confusion matrix heat map, respectively. Figure 32 and 33 depicts the results for Plate Recognizer with CENPARMI dataset and its confusion matrix heat map, respectively.

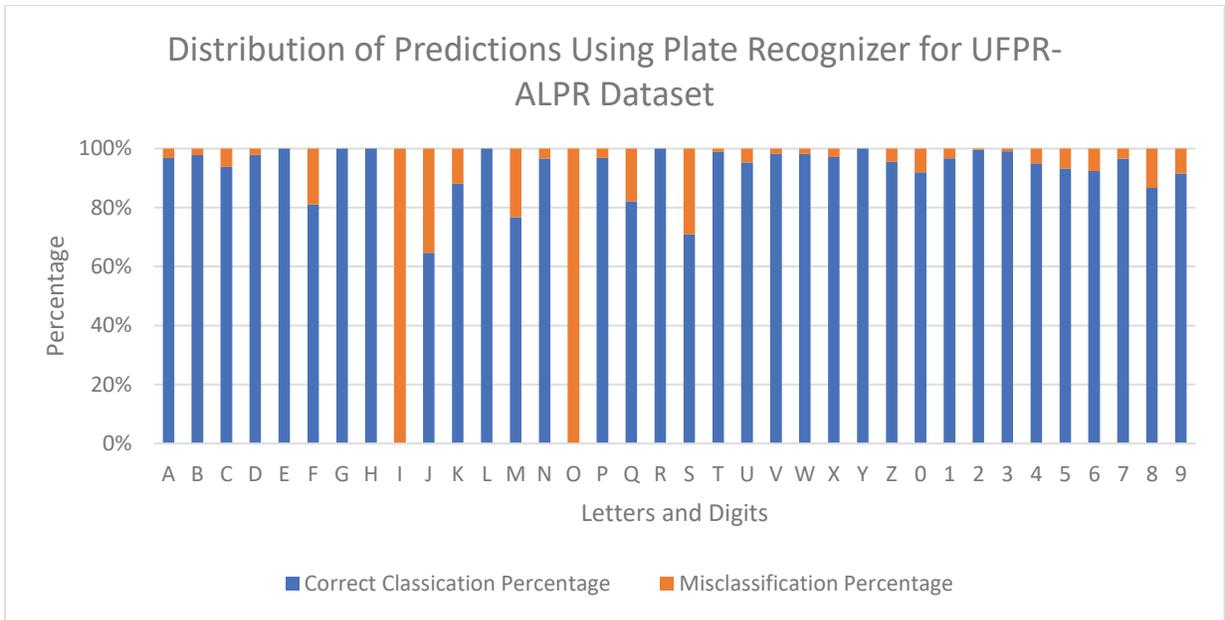


Figure 26. Detection Distribution for all Letters and Digits using Plate Recognizer for UFPR-ALPR Dataset

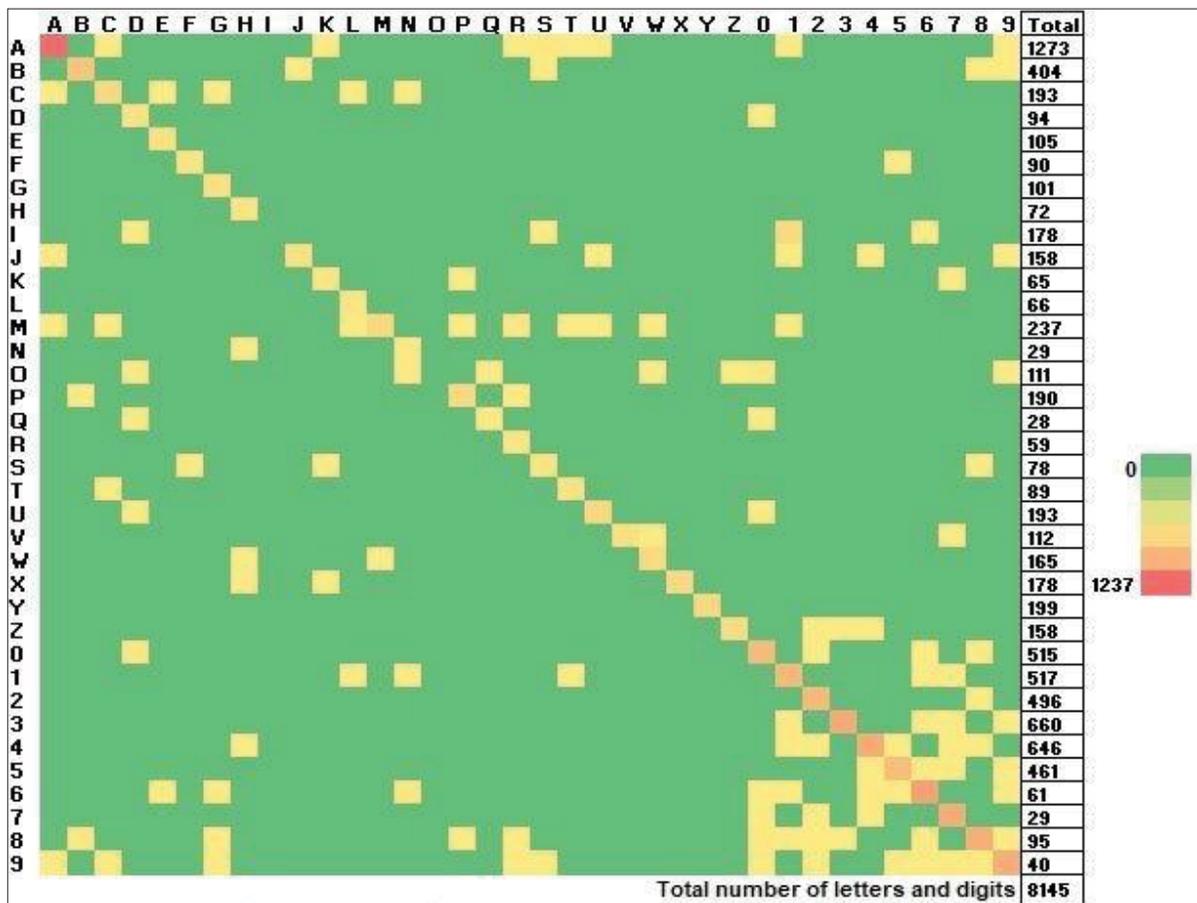


Figure 27. Heat Map of the Confusion Matrix for Plate Recognizer using UFPR-ALPR Dataset

From Figures 26 and 27, we can identify the following problematic cases (severe confusion cases)

- Letter I, fully confused with digit 1.
- Letter O, fully confused with letters D, Q, and digit 0.
- Letter J, confused with A and 1.
- Letter S, confused with F.
- Letter Q, confused with D.
- Digit 8, confused with B and 0.

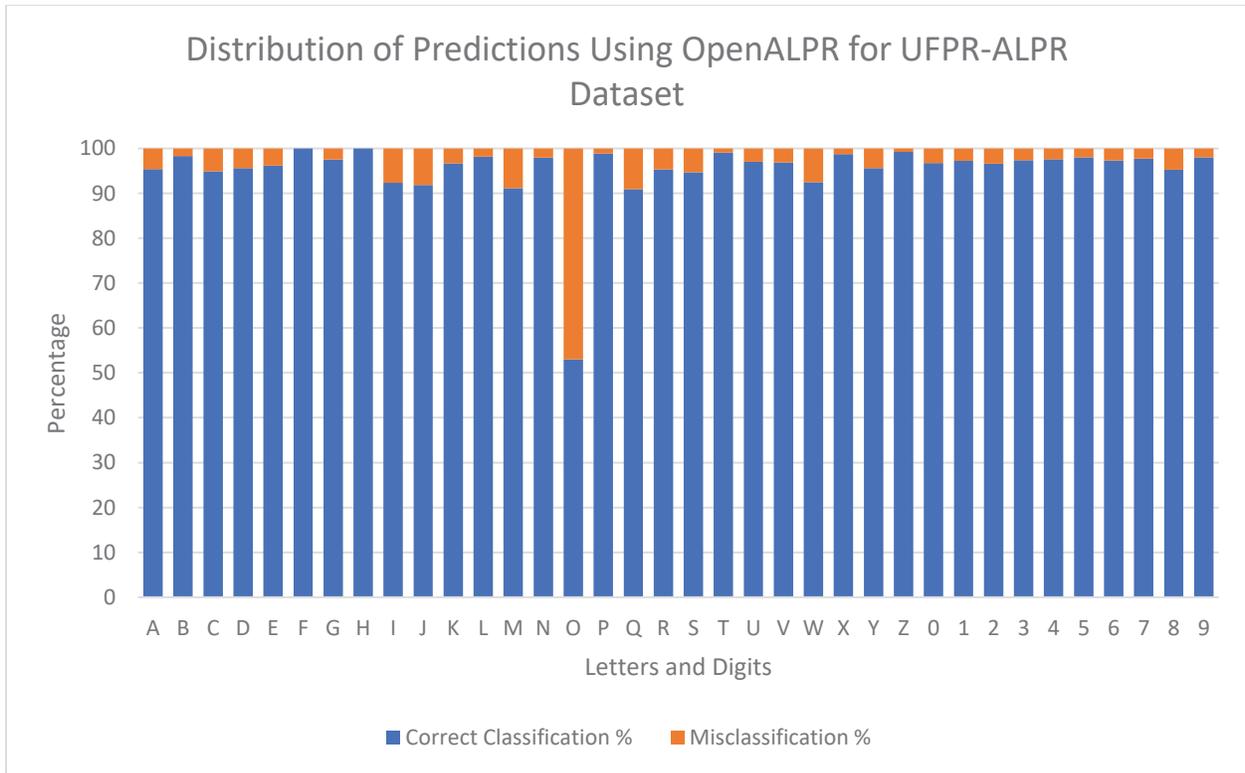


Figure 28. Detection Distribution for all Letters and Digits using OpenALPR for UFPR-ALPR Dataset

From Figures 28 and 29, we can identify the following problematic cases: letter O was confused mainly with both letters D and Q. Table 10 depicts the glyphs classification percentages for UFPR-ALPR using both products.

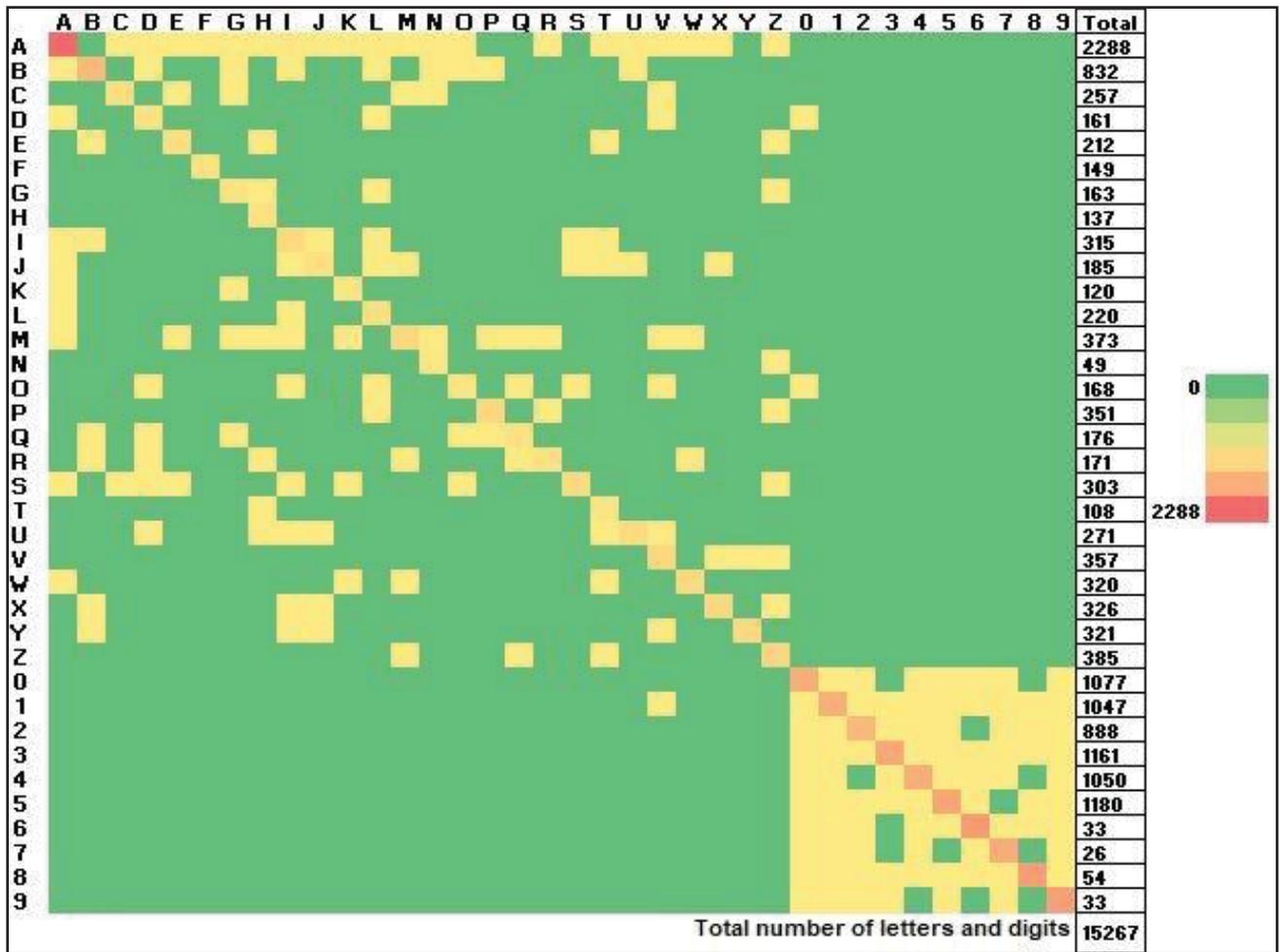


Figure 29. Heat Map of the Confusion Matrix for OpenALPR using UFPR-ALPR Dataset

Table 10. UFPR-ALPR Dataset Classification Percentages using both Products

	OPENALPR		Plate Recognizer	
	Correct Classification %	Misclassification %	Correct Classification %	Misclassification %
<b>A</b>	95.41	4.59	96.79	3.21
<b>B</b>	98.32	1.68	97.78	2.22
<b>C</b>	94.94	5.06	93.78	6.22
<b>D</b>	95.65	4.35	97.87	2.13
<b>E</b>	96.23	3.77	100.00	0.00
<b>F</b>	100.00	0.00	81.11	18.89
<b>G</b>	97.55	2.45	100.00	0.00
<b>H</b>	100.00	0.00	100.00	0.00
<b>I</b>	92.38	7.62	0.00	100.00
<b>J</b>	91.89	8.11	64.63	35.37
<b>K</b>	96.67	3.33	88.06	11.94
<b>L</b>	98.18	1.82	100.00	0.00
<b>M</b>	91.15	8.85	76.79	23.21
<b>N</b>	97.96	2.04	96.55	3.45
<b>O</b>	52.98	47.02	0.00	100.00
<b>P</b>	98.86	1.14	96.84	3.16
<b>Q</b>	90.91	9.09	82.14	17.86
<b>R</b>	95.32	4.68	100.00	0.00
<b>S</b>	94.72	5.28	70.89	29.11
<b>T</b>	99.07	0.93	98.88	1.12
<b>U</b>	97.05	2.95	95.34	4.66
<b>V</b>	96.92	3.08	98.23	1.77
<b>W</b>	92.50	7.50	98.18	1.82
<b>X</b>	98.77	1.23	97.19	2.81
<b>Y</b>	95.64	4.36	100.00	0.00
<b>Z</b>	99.22	0.78	95.57	4.43
<b>0</b>	96.77	3.23	91.95	8.05
<b>1</b>	97.24	2.76	96.74	3.26
<b>2</b>	96.64	3.36	99.60	0.40
<b>3</b>	97.43	2.57	99.09	0.91
<b>4</b>	97.62	2.38	95.03	4.97
<b>5</b>	98.07	1.93	93.29	6.71
<b>6</b>	97.35	2.65	92.42	7.58
<b>7</b>	97.72	2.28	96.60	3.40
<b>8</b>	95.23	4.77	86.72	13.28
<b>9</b>	98.03	1.97	91.50	8.50

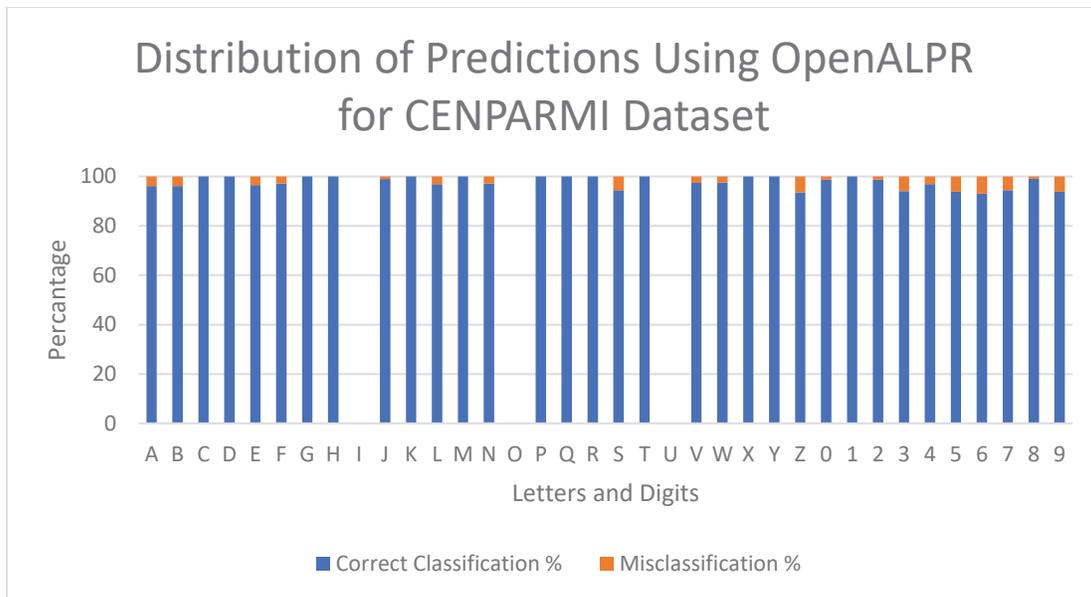


Figure 30. Detection Distribution for all Letters and Digits using OpenALPR for CENPARMI Dataset

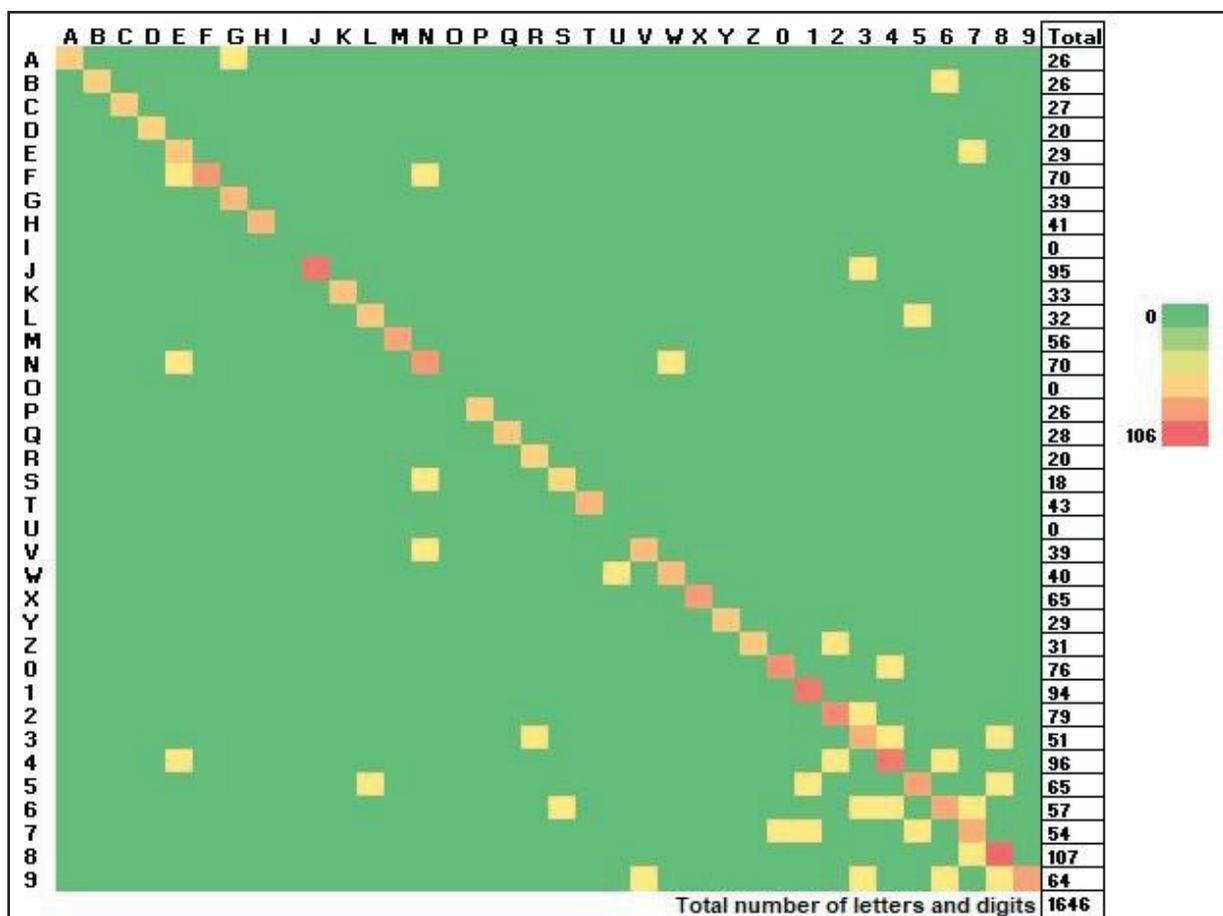


Figure 31. Heat Map of the Confusion Matrix for OpenALPR using CENPARMI Dataset

From Figures 30 and 31, we observe no severe confusion cases are reported. However, we note a slight confusion in the following digits and letters: 5, 6, 9, Z, and S.

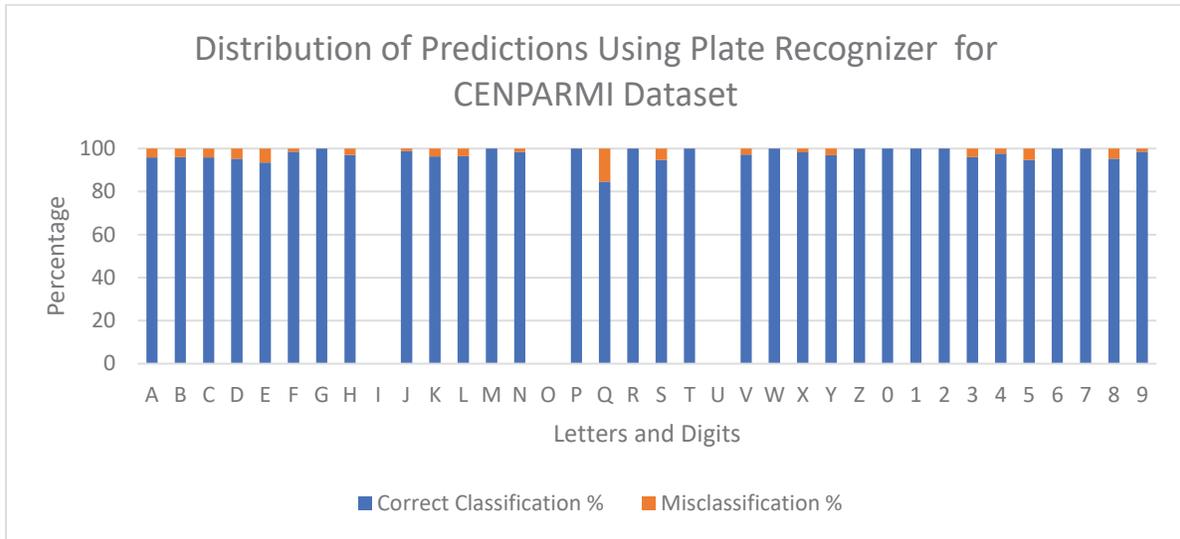


Figure 32. Detection Distribution for all Letters and Digits using Plate Recognizer for CENPARMI Dataset

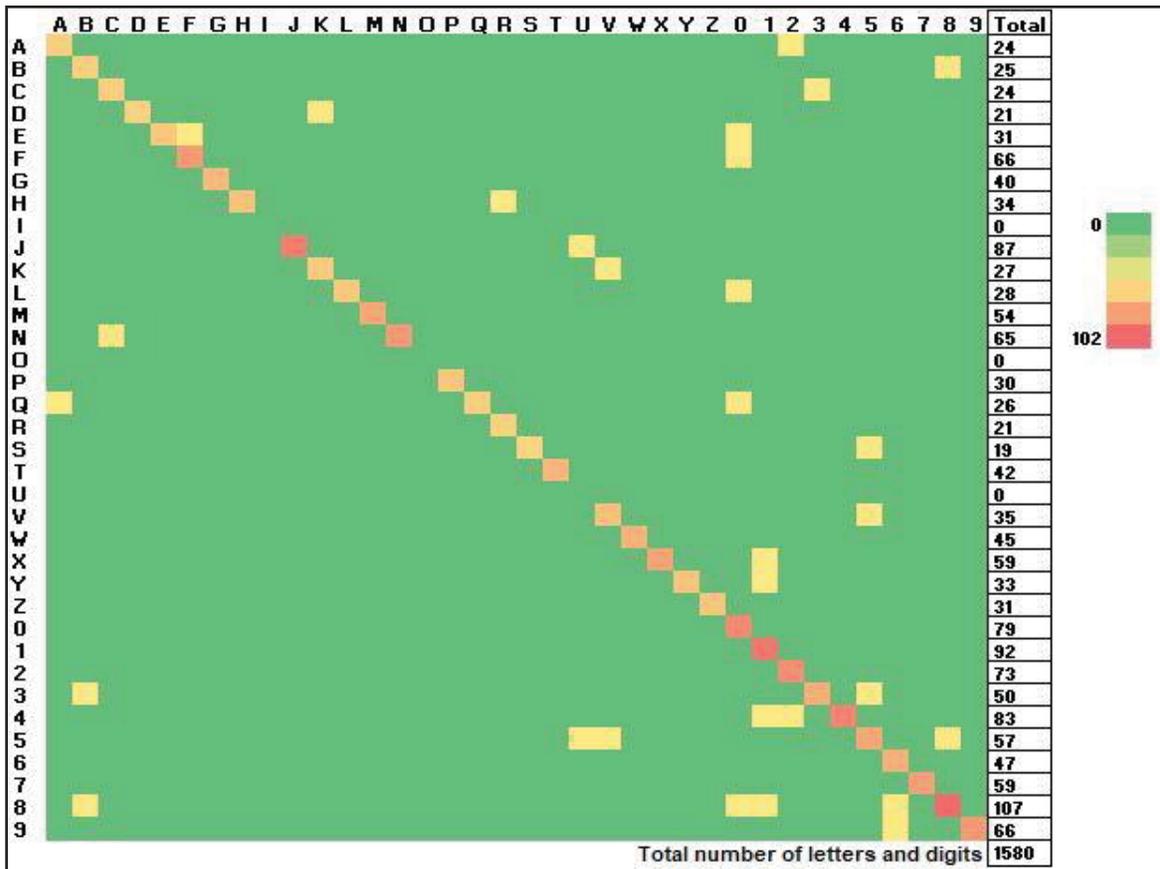


Figure 33. Heat Map of the Confusion Matrix for Plate Recognizer using CENPARMI Dataset

From Figures 32 and 33, we can identify the following problematic cases: letter Q is 15% confused with digit 0. Table 11 depicts the glyphs classification percentages for CENPARMI using both products.

Table 11. CENPARMI Dataset Classification Percentages using both Products

	OPENALPR		Plate Recognizer	
	Correct Classification %	Misclassification %	Correct Classification %	Misclassification %
<b>A</b>	96.15	3.85	95.83	4.17
<b>B</b>	96.15	3.85	96.00	4.00
<b>C</b>	100.00	0.00	95.83	4.17
<b>D</b>	100.00	0.00	95.24	4.76
<b>E</b>	96.55	3.45	93.55	6.45
<b>F</b>	97.14	2.86	98.48	1.52
<b>G</b>	100.00	0.00	100.00	0.00
<b>H</b>	100.00	0.00	97.06	2.94
<b>I</b>	0.00	0.00	0.00	0.00
<b>J</b>	98.95	1.05	98.85	1.15
<b>K</b>	100.00	0.00	96.30	3.70
<b>L</b>	96.88	3.13	96.43	3.57
<b>M</b>	100.00	0.00	100.00	0.00
<b>N</b>	97.14	2.86	98.46	1.54
<b>O</b>	0.00	0.00	0.00	0.00
<b>P</b>	100.00	0.00	100.00	0.00
<b>Q</b>	100.00	0.00	84.62	15.38
<b>R</b>	100.00	0.00	100.00	0.00
<b>S</b>	94.44	5.56	94.74	5.26
<b>T</b>	100.00	0.00	100.00	0.00
<b>U</b>	0.00	0.00	0.00	0.00
<b>V</b>	97.44	2.56	97.14	2.86
<b>W</b>	97.50	2.50	100.00	0.00
<b>X</b>	100.00	0.00	98.31	1.69
<b>Y</b>	100.00	0.00	96.97	3.03
<b>Z</b>	93.55	6.45	100.00	0.00
<b>0</b>	98.68	1.32	100.00	0.00
<b>1</b>	100.00	0.00	100.00	0.00
<b>2</b>	98.73	1.27	100.00	0.00
<b>3</b>	94.12	5.88	96.00	4.00
<b>4</b>	96.88	3.13	97.59	2.41
<b>5</b>	93.85	6.15	94.74	5.26
<b>6</b>	92.98	7.02	100.00	0.00
<b>7</b>	94.44	5.56	100.00	0.00
<b>8</b>	99.07	0.93	95.33	4.67
<b>9</b>	93.75	6.25	98.48	1.52

It is clear that the high confusion cases are in agreement with some of the conclusions drawn from the font anatomy study. Moreover, for the problematic cases identified by the font anatomy study and were not captured by the commercial products, we note that the usage of *regular expression* may have aided these products to escape these cases. However, we expect that in the case of personalised license plates, these products will achieve lower recall ratios.

To improve the glyphs' recognisability, we recommend the following:

1. An open counter for number 4 will lead to straight upper vertical strokes and help to make it look different from the letter A.
2. Different upper and lower counters for letter B help to differentiate the letter B from numeral 8.
3. Designing digit 8 with stroke separation and curvy bowls instead of square bowls to give it a unique shape.
4. Top arm for number 1 helps to differentiate it from letter I.
5. Diagonal straight spine for letter S helps to differentiate it from 5.
6. A straight stem for number 6 will help make it look more different from digit 8.

## Chapter 6. Conclusion and Future Work

This chapter concludes the thesis and provides some recommendations for future work.

### I. Conclusion

License Plate Detection (LPD) is an important topic that is widely investigated by researchers due to its impact in many fields such as traffic monitoring, safety and security. Many techniques were proposed in the literature with the aim of improving the recall ratio. Machine learning and image processing algorithms were the focus of researchers for developing efficient detection algorithms. This work utilizes the emerging DL algorithms to develop new and efficient detection algorithm for complex scenes. However, LDP in complex scenes is a more challenging problem due to its sensitivity to environmental factors (such as rain, dust, and shadow) and lighting conditions, which may greatly influence the detection accuracy. Moreover, LPD is more challenging for real time systems.

Apart from the above, the chosen font type in the license plate plays a vital role in the recognition phase in computer-based studies. Despite its vital importance and great effect on the recognition accuracy, font evaluation for recognition is not sufficiently investigated in the literature. Hence, this work is two folded. On one hand, we utilized Deep Learning (DL) techniques in license plate detection. On the other hand, we evaluated font characteristics in the license plate context. This work examines the effect of customizing important YOLO hyperparameters to improve the detection recall rate for license plate detections. It also performs a qualitative analysis of the false predictions of the trained YOLO models. Furthermore, it analytically evaluates 2 fonts used in license plates and investigate how the typeface design can affect computer-based studies in context.

For the detection problem, a deep learning network was trained, and the detection was achieved in one stage, hence saving computational processing, time and power. The developed deep learning network achieved a detection ratio comparable to other approaches. A 98.3% recall ratio was achieved.

The error analysis conducted in this work revealed insignificant redundancy errors problem. It also showed that the localisation errors are significantly tied to the test resolution, and the irrelevant object errors are tied to the input resolution.

For the font evaluation problem, the work identified key confusion cases between numbers and letters in the used typeface. Moreover, it evaluated the performance of two commercial products in recognizing license plates.

Additionally, this thesis evaluated the performance of 2 commercial license plate recognition products: OpenALPR, and Plate recognizer. The performance was conducted using 2 datasets: UFPR-ALPR, and CENPARMI. The recall ratio results are: OpenALPR with UFPR-ALPR achieved 55.88%, OpenALPR with CENPARMI achieved 51.87%, Plate recognizer with UFPR-ALPR achieved 23.69%, and Plate recognizer with CENPARMI achieved 50.20%.

## **II. Future Work**

The topic of license plate recognition is of vital importance to enhance the level of security in modern societies. This work can be further extended to tackle the detection problem in new and more challenging contexts. We note that license plates come in different sizes, styles, and use different font types. For the license plate detection problem, we propose to train a deep learning network to detect license plates in all kinds of vehicles regardless of license plates type and size.

Moreover, we propose to establish a benchmark/testing framework to evaluate the performance of commercial products in license plates' recognition. The proposed benchmark will support several datasets from around the world, where each dataset uses a specific font type. It will be designed to be easily integrated with new dataset and new commercial products. The new benchmark will develop standard performance metrics for all datasets and products.

Furthermore, we propose to develop a standard benchmark for typographers to evaluate new font types for license plate detection systems. The typographers will provide the system with the font type glyphs and then the system will automatically evaluate its suitability for the detection and recognition model. It will also highlight some possible confusion cases.

Since personalized license plates are now permitted in Quebec as of 2018, we propose to extend the Quebec license plates dataset to include more personalized items, collect a new dataset from other Canadian provinces that allow personalize LPs such as Ontario, Alberta, Manitoba, and Nova

Scotia, retrain the network with new datasets for personalized LPs, and monitor its performance and possible confusion cases.

## References

- [1] J. Redmon, and A. Farhadi, "YOLO9000: Better, faster, stronger," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, Hawaii, USA, 2017, pp. 6517-6525. DOI:10.1109/CVPR.2017.690
- [2] J. Redmon, and A. Farhadi, "YOLO: Real-Time object detection", Retrieved from <https://pjreddie.com/darknet/yolo/> . Last Accessed July, 19, 2018.
- [3] D. Hoiem, Y. Chodpathumwan, and Q. Dai, "Diagnosing error in object detectors", In Computer Vision, ECCV 2012 - 12th European Conference on Computer Vision, Proceedings. PART 3 ed. 2012. pp. 340-353. (Lecture Notes in Computer Science; PART 3). DOI: 10.1007/978-3-642-33712-3\_25
- [4] OpenALPR, <https://platerrecognizer.com/> , last accessed October, 4, 2018.
- [5] Plate Recognizer, <https://www.sighthound.com/products/sighthound-sentry>, last accessed July, 19, 2018.
- [6] Y. Li, and C. Y. Suen, "Typeface personality traits and their design characteristics", In Proceedings of the 9th IAPR International Workshop on Document Analysis Systems (DAS '10). ACM, New York, NY, USA, pp. 231-238, 2010, DOI=<http://dx.doi.org/10.1145/1815330.1815360>
- [7] P. Baines, and A. Haslam, "Type & typography", ISBN-13: 978-0823055289, Watson-Guption, New York, 2005.
- [8] The Oxford Pocket Dictionary of Current English, "License plate", Retrieved from <http://www.encyclopedia.com/humanities/dictionaries-thesauruses-pictures-and-press-releases/license-plate>, Last Accessed September 30, 2018.
- [9] M. Rafique, W. Pedrycz, and M. Jeon, "Vehicle license plate detection using region-based convolutional neural networks", Soft Computing, 22: 6429, 2018, DOI: <https://doi.org/10.1007/s00500-017-2696-2>.
- [10] R. Wang, N. Sang, R. Wang, and X. Kuang, "Novel license plate detection method for complex scenes", 2013 Seventh International Conference on Image and Graphics, Qingdao, 2013, pp. 318-322. DOI: 10.1109/ICIG.2013.69
- [11] W. Kusakunniran, K. Ngamascharyakul, C. Chantaraviwat, K. Janvittayanuchit, and K. Thongkanchorn, "A Thai license plate localization using SVM", 2014 International Computer Science and Engineering Conference (ICSEC), Khon Kaen, 2014, pp. 163-167. DOI: 10.1109/ICSEC.2014.6978188

- [12] Z. Shaikh, U. Khan, M. Rajput, and A. Memon, "Machine learning based number plate detection and recognition", In Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2016), Maria De Marsico, Gabriella Sanniti di Baja, and Ana Fred (Eds.). SCITEPRESS - Science and Technology Publications, Lda, Portugal, 327-333. DOI: <http://dx.doi.org/10.5220/0005750203270333>
- [13] V. Joy, A. Binu, and A. Kuttyamma, "License plate detection and recognition using a HOG feature based SVM classifier", International Journal of Innovative Research in Computer and Communication Engineering, Volume 5, Issue 5, pp. 10115-10119, May 2017, DOI: 10.15680/IJIRCCE.2017.0505235.
- [14] A. Abbasa, and A. Rashid, "Saudi Arabia license plate detection based on ANN and objects analysis", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 13, pp. 3740-3751, 2017.
- [15] N. Astawa, G. Caturbawa, M. Sajayasa, and M. Atmaja, "Detection of license plate using sliding window, histogram of oriented gradient, and support vector machines method", Journal of Physics: Conference Series. Volume 95, Number 1, 2018. Series 953012062. DOI: 10.1088/1742-6596/953/1/012062.
- [16] M. Hall, and L. Smith, "Feature subset selection: a correlation based filter approach", International Conference on Neural Information Processing and Intelligent Information Systems, pp. 855-858). Berlin: Springer, 1997.
- [17] Q. Fu, Y. Shen, and Z. Guo, "License plate detection using deep cascaded convolutional neural networks in complex scenes", In: Liu D., Xie S., Li Y., Zhao D., El-Alfy ES. (eds) Neural Information Processing. ICONIP 2017. Lecture Notes in Computer Science, vol 10635. Springer, Cham, DOI: [https://doi.org/10.1007/978-3-319-70096-0\\_71](https://doi.org/10.1007/978-3-319-70096-0_71)
- [18] H. Li, R. Yang, and X. Chen, "License plate detection using convolutional neural network", 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, 2017, pp. 1736-1740. DOI: 10.1109/CompComm.2017.8322837
- [19] F. Kurpiel, R. Minetto, and B. T. Nassu, "Convolutional neural networks for license plate detection in images", 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 3395-3399. DOI: 10.1109/ICIP.2017.8296912.
- [20] C. Lin, Y. Lin, and W. Liu, "An efficient license plate recognition system using convolution neural networks", 2018 IEEE International Conference on Applied System Invention (ICASI), Chiba, 2018, pp. 224-227. DOI: 10.1109/ICASI.2018.8394573.

- [21] T. Ying, L. Xin, and L. Wanxiang, "License plate detection and localization in complex scenes based on deep learning", 2018 Chinese Control and Decision Conference (CCDC), Shenyang, 2018, pp. 6569-6574. DOI: 10.1109/CCDC.2018.8408285.
- [22] X. Zhang, N. Gu, H. Ye, and C. Lin, "Vehicle license plate detection and recognition using deep neural networks and generative adversarial networks", *Journal of Electronic Imaging*, Volume 27, Number 4, 043056, 2018. DOI: <https://doi.org/10.1117/1.JEI.27.4.043056>.
- [23] G. Hsu, A. Ambikapathi, S. Chung, and C. Su, "Robust license plate detection in the wild", 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Lecce, 2017, pp. 1-6. DOI: 10.1109/AVSS.2017.8078493
- [24] R. Qian, B. Zhang, F. Coenen, and Y. Yue, "Multilingual and skew license plate detection based on extremal regions", 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, 2017, pp. 850-855. DOI: 10.1109/FSKD.2017.8393386.
- [25] R. Laroca, E. Severo, L. Zanlorensi, L. Oliveira, G. Gonçalves, W. Schwartz, and D. Menotti, "A robust real-time automatic license plate recognition based on the YOLO detector", International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 2018, DOI: 10.1109/IJCNN.2018.8489629.
- [26] V. V. Molchanov, B. V. Vishnyakov, Y. V. Vizilter, O. V. Vishnyakova, and V. A. Knyaz, "Pedestrian detection in video surveillance using fully convolutional YOLO neural network", Proceedings Volume 10334, Automated Visual Inspection and Machine Vision II; 103340Q (2017). DOI: <https://doi.org/10.1117/12.2270326>.
- [27] P. Ren, W. Fang, and S. Djahel, "A novel YOLO-Based real-time people counting approach", 2017 International Smart Cities Conference (ISC2), Wuxi, 2017, pp. 1-2. DOI: 10.1109/ISC2.2017.8090864.
- [28] S. Princiya, and V. Saro, "An innovative method for the vehicle number plate recognition", *International Journal of Advanced Scientific Research & Development (IJASRD)*, Vol. 05, Special Issue 01, Ver. II, Mar 2018, pp. 121 – 131.
- [29] H. Saghaei, "Proposal for automatic license and number plate recognition system for vehicle identification", 1st International Conference on New Research Achievements in Electrical and Computer Engineering, IEEE, Tehran, Iran, vol. 1, Issue 1, pp. 1-7, 2016.

- [30] A. Shaji, G. Joy, J. Mathew, O. Patil, and S. Koshy, "Easy license plate detection and recognition", *International Journal of Science and Research (IJSR)*, Volume 7, Issue 3, March 2018, 1674 - 1677, DOI: 10.21275/ART20181077.
- [31] C. Y. Suen, N. Dumont, M. Dyson, Y. Tai, and X. Lu, "Evaluation of fonts for digital publishing and display", 2011 International Conference on Document Analysis and Recognition, Beijing, 2011, pp. 1424-1436. DOI: 10.1109/ICDAR.2011.307.
- [32] N. Hojjati, and B. Muniandy, "The effects of font type and spacing of text for online readability and performance", *Contemporary Educational Technology*, Volume 5, Number 2, pp. 161-174, 2014. DOI: <http://dergipark.gov.tr/cet/issue/25736/271507>.
- [33] K. Larson, and M. Carter, "Chapter 3: Sitka: a collaboration between type design and science", Series on Language Processing, Pattern Recognition, and Intelligent Systems: Digital Fonts and Reading, World Scientific Publishing Co Pte Ltd, edited by: Mary Dyson and Ching Suen, pp. 37-53 (2016), DOI: 10.1142/9789814759540\_0003.
- [34] S. Beier, "Chapter 5: Designing legible fonts for distance reading", Series on Language Processing, Pattern Recognition, and Intelligent Systems: Digital Fonts and Reading, World Scientific Publishing Co Pte Ltd, edited by: Mary Dyson and Ching Suen, pp. 79-93 (2016). DOI: 10.1142/9789814759540\_0005
- [35] M. Almuhajri, and C. Y. Suen, "Chapter 14: Legibility and readability of Arabic fonts on Personal Digital Assistants PDA", Series on Language Processing, Pattern Recognition, and Intelligent Systems: Digital Fonts and Reading, World Scientific Publishing Co Pte Ltd, edited by: Mary Dyson and Ching Suen, pp. 248-265 (2016), DOI: 10.1142/9789814759540\_0014.
- [36] M. Bernard, C. H. Liao, and M. Mills, "The effects of font type and size on the legibility and reading time of online text by older adults", In CHI '01 Extended Abstracts on Human Factors in Computing Systems (CHI EA '01). ACM, New York, NY, USA, 175-176, 2001, DOI: <https://doi.org/10.1145/634067.634173>.
- [37] B. S. Chaparro, A. D. Shaikh, and A. Chaparro, "The legibility of ClearType fonts", *Human Factors and Ergonomics Society Annual Meeting Proceedings*, Volume 50, Number 17, pp. 1829-1832, October 2006. DOI: <https://doi.org/10.1177/154193120605001724>.
- [38] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: unified, real-time object detection", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788, 2016.

[39] CalculQuebec, Retrieved from <https://wiki.calculquebec.ca/w/Helios/en>, last accessed August, 25, 2018.

[40] D. M. Tavares, G. Caurin, and A. Gonzaga, "TESSERACT Ocr: A case study for license plate recognition in Brazil", *Pesquisa e Tecnologia Minerva*, Volume 7, Number 1, pp. 19-26, 2011.