PGTLP: A Dataset for Tunisian License Plate Detection and Recognition

Amir Ismail, Mehri Maroua, Anis Sahbani, Najoua Essoukri Ben Amara

▶ To cite this version:

Amir Ismail, Mehri Maroua, Anis Sahbani, Najoua Essoukri Ben Amara. PGTLP: A Dataset for Tunisian License Plate Detection and Recognition. 18th International Multi-Conference on Systems, Signals & Devices (SSD),, Mar 2021, Monastir, Tunisia. pp.661–666. hal-03910186

HAL Id: hal-03910186
https://hal.sorbonne-universite.fr/hal-03910186v1

Submitted on 3 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PGTLP: A Dataset for Tunisian License Plate Detection and Recognition

Amir Ismail*‡, Maroua Mehri*, Anis Sahbani†‡ and Najoua Essoukri Ben Amara*
*Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, LATIS-Laboratory of Advanced Technology and Intelligent Systems, 4023, Sousse, Tunisie;
†Sorbonne Université, CNRS, Institute for Intelligent Systems and Robotics (ISIR), Paris, France;
‡Enova Robotics, Novation City, Technopôle de Sousse, 4000, Sousse, Tunisie
Emails: amir.ismail@eniso.u-sousse.tn, maroua.mehri@eniso.u-sousse.tn,
anis.sahbani@enovarobotics.com and najoua.benamara@eniso.rnu.tn

Abstract—**The whole point throughout this work is a publicly available and annotated images dataset for vehicle license plate (LP) detection and recognition. This contribution is driving by the fact that no Tunisian LP dataset has been provided up to this point in time. Pearl Guard Tunisian LP (PGTLP) dataset is fully annotated database that contains presently up to $3,000$ images captured by high definition quality camera mounted on a security mobile robot, called Pearl Guard and made accessible to the Academia. For benchmarking and ranking supply, we propose a unified, real-time (LP) detection module based on YOLOv4-tiny detector. We also provide baseline results for vehicle LP detection using the PGTLP dataset.**

Index Terms—**Vehicle license plates, detection, recognition, dataset, benchmark.**

I. Introduction

Governments around the globe are racing to infuse technology into just about every aspect of its city's operations. Many cities start to adopt smart technology to improve their environment and daily living. One of the things that most cities are grappling with their congestion is how transportation works in the city. In an urban environment, where population keeps rising, there is tons of daily traffic. Smart cities nowadays becomes the highest priority including public transportation, E-governance, urban mobility, etc. Given that, the government can effectively monitor its crowd density, cleanliness of public spaces, and even the exact movement of every locally registered vehicle. However, millions of vehicles pass through our cities every day and there is noway to keep track of these vehicles manually.

In this context, license plate detection and recognition (LPDR) is the cornerstone of sustainable urban planning and smart city building. Such a system involves identifying vehicles by their license plate (LP) regions. For vehicles, LP is an important source of information since it carries the distinct identification of the vehicle. In particular, LP recognition is helping cities to manage their traffic flows in an efficient way and it saves time and money to help better manage city parking. Numerous situations that call for LP capture may be very specific, but this feature is incredible important if you need to:

- **Parking management**: A popular application is access control for parking structures (e.g., companies, malls and housing estates). A camera is placed at the entrance gate, and will automatically record an approaching vehicule's LP to compare it to a database of approved visitors. If that LP is on the database, the gate will open and allow the vehiculr to enter based on if the vehicule is on the black list or not. This tends to be a contactless solution during COVID-19.
- **Security**: This will be very useful for law enforcement officers in order to locate a vehicle of interest, search for stolen vehicles or vehicles committing offences. The system can capture LP and match them against a police database list for automatic decisions. For wanted vehicles, an alert informs the relevant authorities for further interventions. Also, it could be extremely useful when it comes to country borders control in order to avoid human error and reduce infrastructure cost. Secure an area, such as a military base or research facility, is as well required.
- **Mobility**: This is particularly important where cities suffer from increasing traffic congestion and its related side effects. LDPR can enable management of the distribution of transportation more efficiently to prevent overflow during traffic peak. One of the most important applications of LPDR in this regard

is statistics collection and hence analysis of traffic during peak periods. It is also a solution for free-flow tolling since it can avoid stop-and-go scenario allowing collection of tolls without interfering with the flow of traffic.

The aforementioned applications are a momentary view to the unlimited use-cases where LPDR systems could potentially and effectively be used. That been said, LPDR systems attracts more and more attention the recent years and countless attempts are made intensively to come out with end-to-end robust and elegant LPDR systems. In practice, LPDR should identify both moving and stationary vehicles in high resolution regardless of lights or weather conditions. Moreover, LPDR approaches should overcome dust, smoke, vehicle speed, etc. Furthermore, LPDR approaches should adjusts to lighting conditions and distance.

LPDR in Tunisia stands out to be not yet mature since not too many systems [1], [2] were presented to address the issue of Tunisian LP recognition. However, up to now and to the best of our knowledge there is no publicly available LP dataset for Tunisian vehicles. This tends to be a critical issue when it comes to develop and test detection and recognition models. Thus, the importance of having a publicly accessible annotated dataset is always recognized by the computer vision research community particularly LPDR researchers.

As a result of this work, we have gathered a dataset of images covering Tunisian LP. It is the first Tunisian database that we make available for studying and benchmarking LPDR approaches. Our dataset, named Pearl Guard Tunisian LP (PGTLP), contains presently up to $3,000$ high resolution images annotated in terms of LP location. What makes this database relevant is that it was collected by an all terrain mobile robot, called Pearl Guard manufactured by our industrial partner "Enova Robotics" for security purposes.

In Section II of this paper, we present an overview of the datasets related to LPDR systems existing in the literature. The content of the PGTLP dataset including details about the ground truth annotations are detailed in Section III. Then, we present a tiny version YOLOv4 detector and we report on its performance on our dataset in section IV. Section V concludes the paper.

II. State-of-the art LPDR datasets

The available LPDR datasets are characterized in particular by the country LP template specifications. Many related works were driven by their particular databases specific for their countries. It is therefore important to highlight the available datasets and their statistics. The following datasets happen to be the most relevant ones in LPDR field.

- **AOLP**: is a public dataset containing $2,049$ images of Taiwanese LP. This dataset is divided into access control (AC), law enforcement (LE), and road patrol (RP) subsets. The AC subset contains 681 images of vehicles passed through fixed passages such as toll stations. A total of 757 images captured by roadside cameras which are used for checking traffic violations are included in LE subset. Lastly, the RP with 611 images is considered the hardest since it contains a lot of samples with oblique LP. Related works used to train on two subsets (AC,LE) while the third subset (RP) was used for testing. For the annotation, there are only the plate bounding boxes (BB) given in the ground-truth folders [3].
- **MediaLab**: consists of 716 images containing Greek LP provided by the multimedia technology laboratory in the national university of Athens. It is divided based on the difficulty level into a normal subset and a difficult one. The difficult subset (D) has 279 images covering situations like shadow, blur and dirt while the simple (S) group contains 437 images.
- **Caltech-cars**: contains 126 images of vehicles from different states of USA. The images have a resolution of 896×592 pixels and they were captured at Caltech parking lot.
- **PKU**: contains $3,828$ images with Chinese LP captured under diverse scenarios. Mainly, it contains five separate groups (G1-G5) corresponding to different configuration environments. The G1, G2 and G3 groups contain only one vehicle instance and consequently one LP. The images in G4 and G5 come with multiple LP [4].
- **KarPlate**: is a Korean car plate database that is divided into three categories. Each subset is intended to be used for a specific task; the LPD is dedicated to the LP detection while the LPR is essentially for the recognition. The third subset named EER is for the end-to-end recognition [5].

Table I recapitulates the characteristics of the existing LPDR databases in terms of image quantity, resolutions and annotations.

Many issues come to the front when dealing with these datasets. In particular, none of them provides bounding boxes for LP detection neither for character recognition. Actually this raises to be a missing feature for these databases especially when it comes to the benchmarking step. Researchers will have to make their appropriate annotations in order to evaluate their models which is very costly and time consuming indeed but also the comparison later on will be unfair since each work will report performance on a completely different ground-truth annotations.

III. PGTLP DATASET

The proposed dataset, named Pearl Guard Tunisian LP (PGTLP), is developed, in cooperation between ”Enova Robotics” and LATIS laboratory to provide the LPDR researchers with a suitable dataset in order to push ahead their research and development of LPDR systems.

A. Image acquisition

The PGTLP dataset was collected using a mobile robot, called Pearl Guard and showed in Figure 1, capable of patrolling various terrains and environments. This robot is equipped with three cameras: optical, thermal and 360 panoramic to be able to almost function whatever the situation. We have used the optical camera to record high quality videos while the robot navigates different environments (e.g., parkings, limited access and high risk areas). It is an AXIS Q1786-LE Network Camera suitable for both indoors and outdoors scenarios. With an outstanding full zoom range up to 32x optical zoom, video in up to 4MP, the camera is almost able to capture any targeted surveillance area with exceptional details for identification and recognition. Moreover, two scenarios were taken into consideration: first when the robot is in motion and the vehicles are parked or both of them are moving. Illustrations of both scenarios are given in Figure 2.

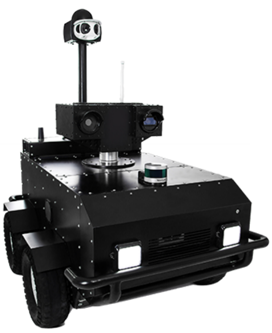

Fig. 1. Pearl Guard robot.

(a) The robot is patrolling and the vehicles are parked.

(b) The robot and the vehicles are moving.

Fig. 2. Illustration of two PGTLP sample images of the rest/mobility scenarios.

B. Characteristics and statistics

Vehicle registration number in Tunisia has different plate formats depending on the field of usage of every vehicle (e.g., rental, ministerial, diplomatic and military). In fact, Tunisian plate particularly combines multilingual characters from either Arabic or Latin languages all along with digits. On the one hand, the standard plates should respect a predefined pattern (cf. Figure 3(b)) that could be segmented into three essential parts: two regions for the digits and ”Tunisia” written in Arabic script situated between them colored in white with a black background. On the other hand, numerous application-specific templates differ fundamentally from the conventional plates. It is also worth pointing out the variety of fore/back-ground colors (white, black, blue and red) and also some plate instances may sometimes contain the Tunisian flag. In the PGTLP dataset, almost every template pattern is covered. Instances, presented in Figure 3, detail the use case of every template.

The PGTLP dataset presently includes up to $3,000$ images gathered from the camera of the mobile robot Pearl Guard. Arguably this was very profitable to ensure maximum diversities of the content in terms of font, size, position and background whereas the robot is pa-

TABLE I
STATISTICS OF THE STATE-OF-THE-ART LPDR DATASETS.

Dataset		Number of images	Image resolution	Annotation		Availability
				Detection	Recognition	
AOLP	AC	681	352 x 240	Yes	Yes	Yes
	LE	757	640 x 480	Yes	Yes	
	RP	611	320 x 240	Yes	Yes	
MediaLab	Simple	437	1792 x 1312 800 x 600 640 x 480	No	No	Yes
	Difficult	279	640 x 480	No	No	
Caltech-cars		126	896 x 592	No	No	Yes
PKU	G1	810	1082 x 728	No	No	Yes
	G2	700		No	No	
	G3	743		No	No	
	G4	572	1600 x 1236	No	No	
	G5	1152	1600 x 1200	No	No	
KarPlate	LPD	4267	1920 x 1080	-	-	No
	LPR	4627		-	-	
	EER	929		-	-	

(a) Temporary registration

(b) Standard Vehicle

(c) Rental Vehicle

(d) Ministerial Vehicle

(e) Diplomatic Vehicle

(f) Military Vehicle

Fig. 3. Multiple templates of Tunisian LP

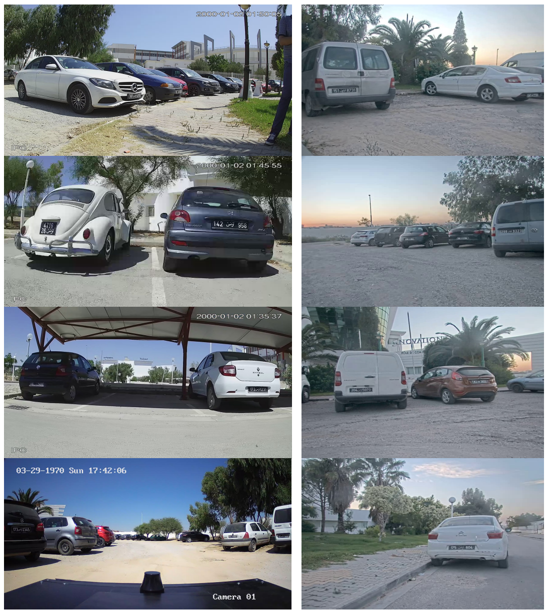

Fig. 4. Samples from the PGTLP dataset. The resolution of the images in the left column is 1920×1080 pixels, while in the right column is 800×600 pixels.

trolling numerous environments without any constraints. Our dataset incorporates three resolutions: 1920×1080, 800×600 and 640×480 pixels. Figure 4 visualizes some images of the dataset. In addition, our dataset considers the case of identifying simultaneously many vehicles. Hence, the data images cover not only one LP per image, but also two even three LP per image.

Typically, any newly introduced database should include annotations and labels for the images. In our case, the ground-truth folder contains one text file per image. Every single annotation file contains in the first column a numeric representation of the label followed by the bounding box annotation. The four values are the center (x,y), the width and the height of the bounding box and they are normalized to lie within the range [0, 1] which makes them easier to work with even after scaling or stretching images. It has become quite popular to follow the Darknet framework's implementations of YOLO detectors. *LabelImg* tool was used to manually label the dataset with bounding boxes.

The obtained dataset is made up of a set of pairs (X, Y), where X is some input example and Y is its associated annotation. It is recommended to split the dataset into three folds, commonly a training, validation and test fold with proportions of 80%, 10% and 10% respectively. The training fold should be used for optimizing the parameters with back-propagation technique [6], the validation fold to determine the hyperparameters of the model, and the test fold to evaluate the model.

IV. EXPERIMENTS AND RESULTS

In order to provide a benchmark/baseline for future evaluation studies in LPDR, a deep model that is the tiny YOLOv4 detector, is assessed using the PGTLP dataset.

A. Evaluated method

The tiny YOLOv4 object detector is the reduced version of the state-of-the-art object detection model YOLOv4 [7]. *YOLOv4-tiny* builds on the progress of YOLOv4, which is the larger full model, but emphasizes model speed and a smaller model size for inference even in small constrained compute environments. This model achieves 40.2% on the MS COCO benchmark [8]. It is significantly less accurate compared to 64.9% with YOLOv4 meanwhile it achieves 371 frame per second (FPS) using GTX 1080Ti way faster than the full YOLOv4 version. With that being said, the tiny version is extremely suitable to identify a single object, such a LP, and for any embedded inference device. To detect our custom LP class, we have to make some adjustments to the model architecture. Therefore, the number of classes and the number of filters in the detection layers were set to 1 and 18 respectively. We have trained the detector using stochastic gradient descent (SGD) algorithm with the momentum of 0.9 and weight decay of 0.0005. We have set the learning rate initially to be of 0.001 and it is decayed by a factor of 10 at the iteration step of $1,600$ and $1,800$. We have selected a maximum number of training iterations equal to $2,000$ and we have used a batch size equal to 64.

B. Performance evaluation metrics

To validate an object detection method, we need to have a way to decide if a giving prediction was correct. Thus, we have computed the intersection over union (IoU) metric between the ground-truth bounding box and the predicted bounding box. The IoU corresponds to the amount of overlap between those two boxes. As we increase the IoU, we require the detector to make up closer prediction to the true value. For example, for $IoU = 0.5$, a predicted LP is considered as true positive (TP) if it has minimum overlap of 0.5 with the ground-truth box. False positives (FP) are the ones with lower overlap. The LP annotated but not detected are considered as missed LP samples and denoted as false negatives (FN). Based on this, precision (P) and recall (R) are defined in equations 1 and 2 respectively.

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$R = \frac{TP}{TP + FN} \tag{2}$$

R is basically a measure of all the positives out there how many of them the model has correctly guessed. It describes how the model guess enough times when there is something to guess. P is another view of thing where it is every time the model guess, did it make a correct prediction. The trade-off is we get less precise by making more predictions at lower confidence score which gives high R. Now, we get into aggregate metrics which summarize the entire P/R curve. The F-score (F) (cf. Equation 3) is a single estimate of the P/R curve and takes both FP and FN into account.

$$F = 2 \times \left(\frac{P \times R}{P + R} \right) \tag{3}$$

Moving towards to the most significant metric which is the mean Average Precision (mAP). Since we are detecting only one class, mAP is actually the same as AP. It looks at 11 various points along the P/R curve and do the average across all those precision values. It is a very effective way to look at results across the entire dataset and avoid biased models in terms of classes or IoU threshold.

C. Results

Experiments have demonstrated that the adopted version of YOLOv4 (*YOLOv4-tiny*) performs very well in detecting the LP region of vehicles. On the precision front, the model can identify the LP and box them in with a precision up to 95.23%. This means that the model detect correctly the LP and it does not get confused with similar-to-LP objects. It can be observed from Table

II that the proposed module reach 97.45% in terms of mAP.

TABLE II
EVALUATION RESULTS ON THE PGTLP TEST SET.

	P (%)	R (%)	mAP (%)	Speed (FPS)
YOLOv4-tiny	95.23	94.21	97.45	90.70

Figure 5 illustrates few result examples of LP detection in PGTLP dataset using the *YOLOv4-tiny* model. By visual inspection of the obtained results, we note that the proposed detector provides satisfying results.

Fig. 5. Result examples of LP detection in *PGTLP* dataset. Images size is 640×480 pixels.

Since the *YOLOv4-tiny* model will be deployed on a mobile robot and operates on video feeds, we care a lot about the computational cost of the module. *YOLOv4-tiny* is capable of running with 90 frames per second (FPS) on NVIDIA Tesla K80 with 12GB of RAM. Another important factor to consider is the memory consumption of the model when it comes to deploying and production. After training and fixing the model parameters, the size of the module is 22MB which makes it suitable for on-edge applications such as the security robot.

V. CONCLUSIONS AND FURTHER WORK

The overall takeaway from this work is an image dataset named *PGTLP* dedicated to the development and evaluation of LP detection and recognition approaches targeting Tunisian vehicles. In summary, we put available the first Tunisian LP dataset containing up to $3,000$ annotated high resolution images captured by the *Pearl Guard* which is a mobile robot of our industrial collaborator *Enova Robotics*. Our dataset covers numerous challenges such as different templates, angles, environment backgrounds making it a major contribution and subject to further experiments and contributions. We evaluated the tiny version of YOLOv4 detection algorithm as proof-of-concept of the presented dataset. Although its diversity, this dataset is still under construction and more images will be added periodically to cover the diversity of Tunisian LP templates.

ACKNOWLEDGMENT

This work has been supported by the VRR research fund from the Tunisian Ministry of Higher Education and Scientific Research that is gratefully acknowledged. The authors would like also to thank our industrial partner ”Enova Robotics” for providing access to their security robot.

REFERENCES

[1] Y. Kessentini, M. D. Besbes, S. Ammar, and A. Chabbouh, “A two-stage deep neural network for multi-norm license plate detection and recognition,” *Expert Systems with Applications*, vol. 136, pp. 159–170, 2019.

[2] Z. Selmi, M. B. Halima, U. Pal, and M. A. Alimi, “Delp-dar system for license plate detection and recognition,” *Pattern Recognition Letters*, vol. 129, pp. 213–223, 2020.

[3] G.-S. Hsu, J.-C. Chen, and Y.-Z. Chung, “Application-oriented license plate recognition,” *IEEE transactions on vehicular technology*, vol. 62, no. 2, pp. 552–561, 2012.

[4] K. Yan, Y. Tian, Y. Wang, W. Zeng, and T. Huang, “Exploiting multi-grain ranking constraints for precisely searching visually-similar vehicles,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 562–570.

[5] C. Henry, S. Y. Ahn, and S.-W. Lee, “Multinational license plate recognition using generalized character sequence detection,” *IEEE Access*, vol. 8, pp. 35 185–35 199, 2020.

[6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[7] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “Yolov4: Optimal speed and accuracy of object detection,” *arXiv preprint arXiv:2004.10934*, 2020.

[8] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, “Microsoft coco: Common objects in context,” in *European conference on computer vision*. Springer, 2014, pp. 740–755.