



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 9, Issue 9, September 2020

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Complete Automated License Plate Recognition Using Deep Learning Techniques - A State of The Art Review

P.Banumathi¹, J.C. Miraclin Joyce Pamila²

P.G. Student, Department of Computer Engineering, Government College of Technology, Coimbatore, India¹

Associate Professor, Department of Computer Engineering, College of Technology, Coimbatore, India²

ABSTRACT: Automatic license plate detection is one of the core tasks in numerous applications such as toll management, parking management, traffic enforcement, stolen vehicle detection, road safety management, and many more. Owing to the variety of plate characteristics and environmental factors during camera calibration, ALPR poses a challenge. Numerous ALPR methods are available based on deep learning and computer vision, and these systems have used different kinds of methodologies. The vast majority of methodologies perform under confined conditions such as fixed illumination, stagnant backgrounds, country-wide, specific to particular vehicle and position, and so on. ALPR procedures are usually divided into four steps with vehicle identification, license plate identification, segmentation of characters, and character recognition. The performance of the vehicle detection and license plate detection stages decides the efficiency of character recognized at the final stage. In this article, ANPR techniques are addressed and briefed with factors such as accuracy rate, processing time, plate characteristics, and so forth. So this study will assist the recent researchers in the ALPR application areas to distinguish advanced techniques and their performance. Eventually providing a comparative outline that will provide brief description of all the methods with their capabilities and limitations.

KEYWORDS: License Plate Recognition, Convolutional Neural network, OCR, object detection.

I. INTRODUCTION

ALPR is an intelligent system related to transportation and it is growing day by day. Several works have contributed to the emergence of many such algorithms related to image processing, machine intelligence, and computer vision. Most ALPR systems are built for both industrial and academic use. Recently, the researchers also implemented various algorithms based on state-of-the-art profound neural networks of convolution. The intention of vehicle license plate identification seems to be to locate the license plate attached to a vehicle and then detach the characters from the license plate. A registration number is widely considered the registration plate /number plate of a vehicle. ALPR generally classified into four stages: object detection, license plate detection, character segmentation, and recognition. Detection of artifacts requires recognizing vehicles seen in a picture. License plate identification is the way to locate a license plate anywhere in a detected vehicle. Character segmentation finds sequence characters in the image and subdivides into each character region. Character recognition matches each region's character into the exact one. Past few year CNN based techniques developed for object detection, classification, and recognition.

Recently, the "You Just Look Once" (YOLO) object detection network [15, 10, 5] prompted the most researchers and continued to build networks relying on the Yolo network [2, 4, 6-9].YOLO handles object detection and classification inside one network. Nowadays privately owned vehicles are increased day by day and multiple objects in the single image need be detected, YOLO solves these issues by detecting multiple objects in the single pass. The task of license plate detection can broadly be categorized into the following three approaches [4]; 1) Regional, 2) Pixel manipulation, and 3) Focusing Color. In region-based methods [4, 11, 12] input image is segmented into smaller regions and features are extracted to find license plates. To identify LP, each pixel is evaluated in pixel to pixel approach to reach each of its neighboring pixels. In the Color-based approach, the RGB images converted into HSV space and using filters to find the LP regions. Several of the applications incorporate temporary redundancy [6, 14] to positively diagnose the characters on a license plate. Deep learning object detectors classified into two one stage and two-stage detectors [14]. Two-stage detectors result in the formation of candidate regions by an RPN network, one stage network predicts scores for the default set of bounding boxes and eliminates the need for regions. First two-stage network R-CNN [21] for object detection with CNN three-stage regional proposal, feature extraction and

classification, Later R-CNN [20] improved to fast the training and evaluation time by all three stages in a single network. Later Faster R-CNN [12] introduce multi-scale anchors to handle scale invariance.

Convolutional Neural Network (CNN) is a layer of convolutions and is generally used for image recognition, labeling, segmentation, and other auto correlated data as well. A convolution primarily transfers data over the filter. CNN's are ANNs where the idea of full connectivity is substituted by filter convolutions. So for every framework that has structured inputs that can be processed using convolutions. The network can process frame recordings or pictures. It may also be used for developments, focused on audio signals.

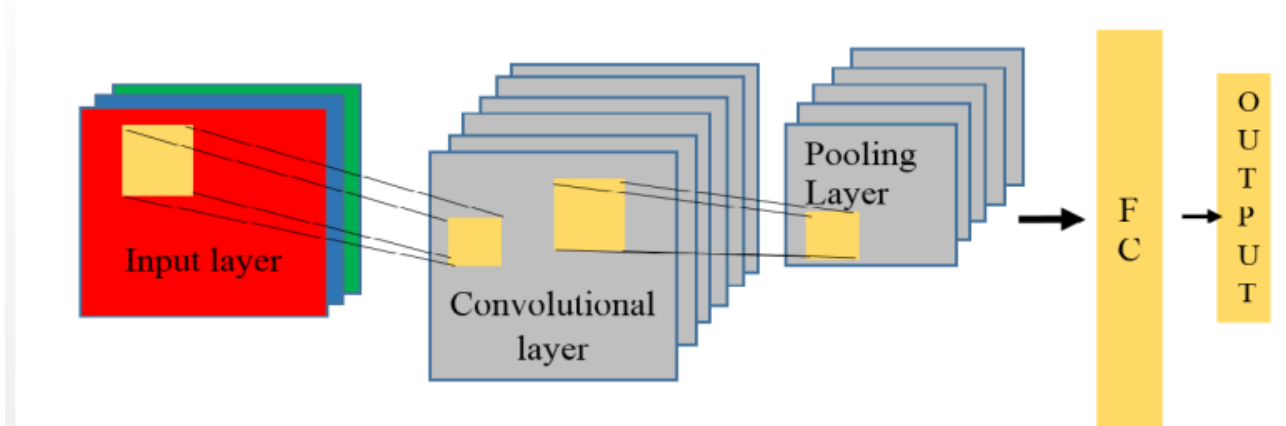


Fig 1. Architecture of general convolutional neural network

Three important parameters of the convolution neural network is the filter size, stride and padding. In convolutional layer output size is smaller than the input image. In order to retaining the size of the image padding is necessary. Stride is about how many shifts a filter move to the right side with related to pixels in the image. The image will pass as input into the CNN layer for feature extraction, known as forward propagation. Follow by the backward propagation is to minimize the error. Rectified Linear Unit (ReLU) layer will perform element-wise activation. The activation function is to increase the non-linearity. ReLU transforms the input so that it is separable. The advantages of ReLU layer is that it works well without pre-training and can run very quick and accurate compare to other function. The pooling layer then progressively reduces the dimensionality of the image to control over fitting. However, to train the CNN with good accuracy and high recognition rate, high amount of training labeled data is needed as a CNN consists of multiple layers of neurons. The CNN mostly have 3x3 filter size [17] and number of filters based on the output feature map and if the feature map size is reduced and the number of filters are increased to preserve layer complexity.

Challenging factors to ALPR: [26]

- 1) Image Resolution: low-resolution images can cause by distance from cameras and low-quality cameras.
 - 2) Blurry shot: this causes due to capturing images during motion, camera’s low shutter speed of the sensors, camera lens debris, and focus depth, and more.
 - 3) External factors: Illumination intensity, shadows, and reflections of light, etc.
 - 4) Internal factors: Distorted textures and sometimes the debris on the plates make elusive.
- Number plates are standardized in some of the developed countries such as Australia, Canada, unites states and the United Kingdom, and so on. In developing countries even though standardized but not fully evolved so they have different plate characteristics than standardized ones. So ALPR more challenges in developing countries. in this paper, objects are referred to as artifacts.

Plate characteristics: [23]

Location: the LP exists in different locations of an image and LP exists mostly bottom or middle of the vehicle. Objects in the image: An image may contain no or many vehicles with license plates. Size: Due to different camera distance and rotation angle, the plates can be of different sizes. Color: Various font color and plate background colors. Fonts: languages and dialects based on the countries. Aspect ratio: The different countries have different plate size.

Environmental characteristics:

The dirt may obscure the plates and distort them, Plates are inclined. Plates are illuminated due to environmental weather, shadow, and intensity. Plates become transparent unless there are other vehicles behind them.

Almost all of the deep learning, computer vision, and machine intelligence advancements indeed recently made significantly better performance, primarily due to data sets availability. In addition, this survey provides a variety of resources used by different authors and then describes its essence as well. This helps the researchers to examine whether the types of resources are suitable for their analysis.

Rest of the paper is organized as follows Section II provides an overview of object detection algorithms. Section III provides an overview of license plate detection methodologies proposed by the various researchers. Section IV provides an overview of the character segmentation and recognition techniques available. Section V provides existing complete ALPR procedures and comparative outline that will provide a brief description of all the methods with their capabilities and limitations. Section VI concludes with a future scope of research.

II. OVERVIEW OF OBJECT DETECTION METHODS

Yolo object detection network recently inspired by many researchers because of its performance and it can perform object detection and classification in one network. Yolo v2 [15] developed in 2016, it is the emerging network for real time object detection that uses a model with 19 convolution layer, 5 max pool layer. And later in 2017 Fast Yolo [10] developed focused on a speed/accuracy trade off, it can classify 9000 objects so it is named also Yolo 9000. It uses 9 convolution layer and less filters in those layers. In 2018 the incremental approach [5] considering both speed and accuracy. It was named Yolo v3, and it contains 59 convolution layers, named Darknet-59. YOLO v3 more accurate than Yolo V2 and fast YOLO.

Basically, YOLO is a single convolution network which simultaneously finds multiple bounding boxes and class probabilities for those boxes. YOLO 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: $(x, y, h, w, \text{confidence})$ where, x and y are the coordinates of the center of the bounding box relative to the grid cell location, h and w are the height and the width of the bounding box relative to the whole image, If no object exists in that cell, the confidence score is zero. Otherwise confidence is the Intersection over Union (IOU) between the predicted box and any ground truth box. Confidence is defined as $\text{Pr}(\text{Object}) * \text{IOU}(\text{truth}, \text{pred})$. The value of $\text{Pr}(\text{Object})$ is 1 when a grid cell contains a part of a ground truth box, otherwise, it is 0. Each grid cell also predicts C conditional class probabilities. YOLO predicts multiple bounding boxes per grid cell.

The authors in [12] proposed object detection using Regional proposal network to find object locations effectively, the system called faster R-CNN. The first stage RPN finds region and share it to the next stage R-CNN. In RPN sliding window 3×3 sliding window low dimension feature is fed into two sibling fully connected layers—a box-regression layer and a box-classification layer on the models Zeiler and Fergus model and Simonyan (ZF, 256-dimension) and Zisserman mode (VGG-16, 512-dimension). In regression layer $4K$ outputs where k is the maximum number of proposal for each region. In classification layer $2K$ scores that predict object probability for each region. Also uses translation invariant which modify the model size if the object translated. And also employs two different ways to share RPN and R-CNN. One is Alternative training in which first trains RPN and region proposal used by Fast R-CNN. Another way is joint training RPN and Fast R-CNN networks are merged into one network during training with appropriate and non-appropriate training that focus ROI pooling layer that is differentiable with respect to the box coordinates.

The author's in [16] proposed a single shot multi detector, it detects multiple objects within the image using single network with 300×300 and 512×512 input size. SSD needs only input image and ground truth box for each object during training. Experimental results conducted on the PASCAL VOC, COCO, and ILSVRC datasets. SSD 300 and SSD 512 are implemented both outperforms R-CNN and YOLO V2 in terms of speed and accuracy. The author's in [24] presented a 3D object representation using fine grained categorization using collected dataset. The basis of 3D object representations is given by estimating the 3D geometry of an object. Providing a frame of reference for both our 3D appearance representation (Patch sampling, Patch rectification and descriptors) and our 3D spatial pooling (3D Spatial Pyramid (SPM-3D), 3D Bubble Bank (BB-3D)), first stage in our fine-grained categorization consists of

identifying one that best fit the image. The matching between a 3D CAD model and a 2D image is implemented by a set of classifiers that have been trained to distinguish among CAD models and viewpoints.

The author's in [17] introduce deep residual learning which are attracted by recent researchers [7]. It reformulates the layers of deep learning neural network as learning residual function with reference to the layer inputs, instead of learning unreferenced functions. ResNets inserts shortcut connections which turn the network into its residual version. Shortcut connections are utilized when the input and output resolutions are same. When the dimensions are increases shortcut performs identity mapping with extra zero padding for increasing dimensions and project shortcut used to match dimensions when the shortcut go across feature maps of two sizes with stride of 2. The model evaluated using on the ImageNet 2012 classification dataset and CIFAR-10 dataset, PASCAL VOC and MS COCO.

III. OVERVIEW OF LICENSE PLATE DETECTION METHODS

The authors in [3] proposed Yolo inspired framework for detecting license plates directly in the complex scenarios without object classification. Model generator able to generate n number of models each trained with single network and testing configurator to generate configurations based on the hyper parameters from the learning phase. The model selector selects best model based on the recall ratio. It states that increasing the training network input resolution or testing network resolution does not necessarily warrant an increase in the recall ratio's and also increasing the training dataset size beyond a certain threshold may not provide increase in recall ratio's.

The author's in [5] proposed a YOLO inspired network called MD-YOLO (multi-directional) which can determine the angle of rotation of a given car license plate image and to compute .IOU (Intersection over Union) using angle deviation penalty factor(ADPF). MD-YOLO executes direct regression to detect a car license plate. MD-YOLO uses 7 Convolutional layer and 3 fully connected layer, 3×3 kernel size and padding size 1. Leaky and identity functions are chosen as the activation function rather than ReLU to identify negative rotation angle. Drawback is two model prepositive (to detect small input size) and MD-YOLO which are trained separately and challenging tasks in illumination, occlusion and so on still remain. The authors in [10] proposed license plate detection based on the Yolo v2 by modifying grid size and bounding box configurations.

License plate detection is based on the CNN [12] and is a type of connection based on the animal visual system implemented by a sequence of layers that perform convolutions. Region based networks is used the images divided into sub regions and value of the sub regions calculated independently if the value is 1 then sub region contains a nearly central LP and 0 when it contains less than half of a LP, with other situations producing an intermediate values. Through estimating the outputs of nearest sub regions it is possible to locate the license plates.

The author's in [26] proposed Edge clustering using expectation minimization (EM) algorithm for plate detection and character segmentation. Maximally stable extreme region (MSER) detector can efficiently identify local regions almost invariant to illumination variation and classifying with null class for non-character region for improving recognition rate.

IV. OVERVIEW OF CHARACTER SEGEMENTATION AND RECOGNITION

The author in [3] presented a Novel Instance Transformation network (ITN) to learn geometric configurations and to detect words or text lines on the scene at one pass. ITN is a multitasking learning with regression transformation on coordinates and text /non-text classifications. ITN also used to detect multi-scale, multi-oriented and multi-lingual scene text recognition. The authors in [13] proposed a fully convolution regression network to train synthetic images for detecting text. Which efficiently performs text detection and bounding box regression at all locations and multiple scales in an image and achieves 84.2% F-measure on the standard ICDAR 2013 benchmark dataset? The network architecture based on VGG-16 and 9 convolution layer. It eases the training and improves efficiency compared to YOLO V2.

The author in [19] introduce spatial transformer network it allows spatial manipulation of data within the network. The spatial transformer mechanism is split into three parts, first a localization network takes the input feature map, and through a number of hidden layers outputs the parameters of the spatial transformation (translation, rotation and more) that should be applied to the feature map –this gives a transformation conditional on the input. Then, the predicted transformation parameters are used to create a sampling grid, which is a set of points where the input map should be sampled to produce the transformed output. This is done by the grid generator. Finally, the feature map and the

sampling grid are taken as inputs to the sampler, producing the output map sampled from the input at the grid points. The combination of the localization network, grid generator, and sampler form a spatial transform network. MNIST handwriting dataset trained on baseline fully-connected (FCN) and convolution (CNN) neural networks, as well as networks with spatial transformers acting on the input before the classification network (ST-FCN and ST-CNN).

The author's in [22] proposed framework for detecting text in the natural scenes without human labeled data. This framework uses three models 90k-way dictionary encoding which describes a model performing pure word classification to a large dictionary, explicitly modeling the entire known language, character sequence encoding that encodes the character at each position in the word, making no language assumptions to naively predict the sequence of characters in an image., and bag-of-N-grams encoding that encodes a word and gives composition model of words with bag-of-N-grams. These models read words in different ways. The datasets used are ICDAR 2003, ICDAR 2013, Street View Text, and IIIT5k. The author's in [18] provides effective license plate character segmentation based on the Jaccard centroid coefficient with binarization and fully connected components techniques. The dataset uses were SSIG-SegPlate Brazilian LP'S.

V. DATASET DESCRIPTION

The authors in [2] used UFPR-ALPR dataset for testing the performance of solutions. This work introduces a new annotated dataset for Canadian license plates (LP), namely CENPARMI datasets. Centre for Pattern Recognition and Machine Intelligence (CENPARMI) dataset of license plates in complex scenes: This dataset consists of 480 images of s from Quebec/Canada. This dataset was captured with an iPhone 6 camera. The images were captured in Bright sunlight, and shade. Distance from camera (1 to 30 meters), Inclined LP, Different angles of Motorcycles, cars and public transportation vehicles. Federal University of Parana dataset for Automatic License Plate Recognition (UFPR-ALPR) [6]: This dataset consists of 4500 images of LPs from Parana /Brazil. This dataset was captured with three different cameras; GoPro Hero4 Silver, Huawei P9 Lite and iPhone 7 Plus. The dataset is split by UFPR as follows: 40% for training, 20% for validation and 40% for testing. The authors in [3] used two benchmark datasets ICDAR2015 and MSRA-TD500. ICDAR2015: It is an incidental scene text dataset which contains 1000 training images and 500 test images. This dataset is very challenging because the images are collected by wearable cameras, so texts in the images appear in random scale, orientation, location, viewpoint and blurring. The annotations of object bounding-boxes represented by 8 coordinates of four clock-wise corners. MSRA-TD500: It contains 300 training images and 200 test images of multi-oriented texts. It is a multi-lingual dataset including English and Chinese. The annotations of MSRA-TD500 are at line levels which are represented by aligned horizontal rectangles and their orientations.

The author in [7] uses ALPR and proposes a new dataset of Brazilian vehicles called UFPR-ALPR. The dataset contains 4,500 images in Portable Network Graphics (PNG) format with size of $1,920 \times 1,080$ pixels. The cameras used were: GoPro Hero4 Silver, Huawei P9 Lite and iPhone 7 Plus. "900 of cars with gray LP, 300 of cars with red LP and 300 of motorcycles with gray LP". It is noteworthy to mention that SSIG Database (SSIG) [18] is the largest public dataset of Brazilian LPs with 2000 images. It uses a static camera mounted always in the same position, all images have very similar and relatively simple backgrounds, and there are no motorcycles and only a few cases where the LPs are not well aligned.

The author's in [4] used AOLP dataset which contains 2049 images: Access Control (AC) 681 samples, Traffic Law enforcement (LE) 757 samples and Road Patrol (RP) 611 samples. The authors in [7] used SSIG Dataset, OPENALPR (BR and EU) and AOLP dataset (LE and RP). The authors in [9] provided an "Application Oriented License Plate Extended (AOLPE) database" which can contains plates undergo weather condition, lightings, complex scenes and so on. The authors in [12] Microsoft COCO object detection dataset with 80 object categories and PASCAL VOC dataset both 2007 and 2012. The dataset in [13] includes scene texts namely SynthText in the Wild(800,000 training images), ICDAR Datasets(IC011, IC013) they contain real world images of text on sign boards, books, posters and other objects with world-level axis-aligned bounding box annotations, Street View Text: It Consists of images harvested from Google Street View annotated with word-level axis-aligned bounding boxes. SVT is more challenging than the ICDAR data as it contains smaller and lower resolution text. The author's in [14] gives new dataset called SSIG-ALPR (Brazilian), contains 6,660 images with 8,683 license plates from 815 different on-track vehicles. However, 3,368 license plates have no text annotation, low-resolution and it is used to labels for LP detection. All images are available in the Portable Network Graphics (PNG) format with size of $1,920 \times 1,080$ pixels. The training set contains 3,595 images, the validation set has 705 and the test contains 2,360 images. The training images are further increase to 800000 samples with data augmentation process such as translation, rotation, and zoom in, zoom out.



VI COMPLETE ALPR SYSTEMS AND COMPARATIVE OUTLINE

Many of the authors proposed only a part of the ALPR system (vehicle detection or LP detection or character recognition). The authors in [6] present a complete ALPR based on the Yolo framework with temporal redundancy that process each frame independently and then combine the results to create a more robust prediction for each vehicle. This outperforms both commercial systems Sighthound and Open ALPR. And the author’s also proposed a new dataset called UFPR-ALPR.

Authors in [8] developed a CNN based complete ALPR for car dataset it can only detect frontal view of the car’s and then detect LP’s finally segment the characters in the LP using modified YOLO v2. In which car frontal view and LP detected in a single network. They consider two performance improvement concepts first to keep only seven characters even detected more than 7 characters using probability. Second the first three characters are assumed to be letters, and the following four digits, these assumption makes even better performance when swapping applied by the confusion matrix that obtained when training. ALPR in this system achieved high accuracy for both frontal view LP detection and character recognition. But the system not applicable for other countries with more than seven characters and it limits to detect only frontal view LP’s.

Then same author’s implemented [7] three stage network Yolo v2 based object detection and classification network, and implemented novel CNN called WPOD-NET insights from YOLO, SSD and Spatial Transformer Networks (STN). YOLO and SSD perform fast multiple object detection and recognition at once, but they do not take spatial transformations into account, generating only rectangular bounding boxes for every detection. STN can be used for detecting non-rectangular regions, however it cannot handle multiple transformations at the same time, performing only a single spatial transformation over the entire input. 21 convolution layers, where 14 are inside residual blocks and 4 max pooling layers. To extract the LP from the detected vehicle fix square around the center of a LP. OCR is performed based on the implementation [8] and complete ALPR runs in 5 FPS it is very low compared to other’s but stated that to increase FPS we can increase vehicle detection threshold but it reduce recall ratio. It is improved to detect both frontal and rear views of various countries (Europe, United States and Brazil) and still the system lacks to detect other vehicles than car LP’s.

The authors in [14] proposed a two deep networks first perform license plate detection directly on the input image and second performs license plate recognition with IOU=0.7 for LP’s. It improves the robustness of network by learning a joint representation that is useful for performing more than one task in same image. In this each task is classified one character in the plate. They have used two stages that directly find LP’s without detecting vehicle first. Limitation is that they have implemented only for Brazilian LP’s.

The author’s in [1] presents complete ALPR based on the YOLO with layout classifications for American, Brazilian, Chinese, European and Taiwanese. Vehicle detection based on the YOLO V2 [15], LP detection and layout classification based on FAST-YOLO [10] used in [8] and LP recognition based on the CR-NET. In this the number of vehicles increased within the same image the FPS is reduced with 5 vehicles per image with 29 FPS.

Table 1. Complete ALPR outline

Paper	Methods	Dataset name and size	Image characteristics	Accuracy rate	Processing rate
[6]	FAST YOLO (object detection and LP detection) but this were not obtained good results for UFPR-ALPR. YOLO V2(object detection and LP detection) OCR network [7] for character segmentation and recognition with temporal redundancy.	SSIG-SegPlate [17] (cars only) (http://www.ssig.dcc.ufmg.br/) 2,000 frame from 101 vehicle videos. 1920×1080 pixels. UFPR-ALPR(cars, buses, motorcycle and trucks) [5] (https://web.inf.ufpr.br/vri/databases/ufpr-alpr/) 45000 frame from 150 vehicle videos with 1,920 × 1,080 pixels.	Static digital camera and fixed background. Captured in complex backgrounds and under different lighting conditions. Capturing devices are GoPro Hero4 Silver, Huawei P9 Lite and iPhone 7 Plus.	Recognition rate of 93.53% for UFPR-ALPR with redundancy	47 Frames Per Second (FPS) for UFPR-ALPR with redundancy



[7]	Vehicle detection using YOLO V2 and LP detection using CNN proposed WPOD-NET (21 convolutional layers and 14 are inside residual blocks [16] using YOLO[14], SSD [15], and STN [18]. The OCR performed using YOLO V2 based network [7].	Annotated data with 196 car images. 105 from the Cars Dataset [23], 40 from the SSIG Database [17], and 51 from the AOLP (LE) dataset [24]. (http://www.inf.ufrgs.br/~crjung/alpr-datasets) For testing OpenALPR (BR and EU),SSIG(test) and AOLP (RP) (https://github.com/openalpr/benchmarks)	Mostly European LPs and others such as USA, Brazilian and Taiwanese LPs. Augmentation used to increase the training dataset.	Average rate of 89.33% for all test sets.	Average of 5 FPS (considering all datasets). On 12Gb of RAM and an NVIDIA Titan X GPU.
[8]	FAST_YOLO to detect car frontal views and LP's by reducing filters for 2 classes. For character segmentation and recognition based on YOLO V2 modified to fit all the LP characteristics	Brazilian SSIG Database [17]2,000 high definition pictures from 101 different cars.	Static camera and fixed car and considering only car data's frontal view.	63.18% accuracy rate for all 7 letters.	76 FPS for high end GPU and 9 FPS for mid end GPU.
[14]	Detect the license plates directly from the image frame using CNN with 10 convolutional layer and character segmentation and recognition using CNN with 7 convolutional layers.	SSIG-ALPR 6,660 images with 8,683 license plates of Brazil.	815 different on-track vehicles. 3,368 license plates with impossible recognition. Obtained by two different camera's static and vehicle moving camera.	88.8% recognition rate in overall.	30 FPS even 6 vehicles in an image which is applicable in real-world situations.
[1]	Vehicle detection using YOLO V2, LP detection and layout classification based on the modified Fast YOLO V2 [8] and LP recognition using CR-NET.	Caltech cars, UCSD-Stills, ChineseLP, AOLP,OPENALPR(EU) , SSIG-SegPlate and UFPR-ALPR 6239 pictures	8 different datasets from US, Chinese, Europe, Brazilian, Taiwanese. Manually labelled 38351 Bounding boxes on images.	96.8% Recognition rate for 8 pubic datasets.	73 FPS in high end GPU and reduced if the number of images increases/ image

VII.CONCLUSION

In this survey, Existing ANPR techniques are addressed and briefed with factors such as accuracy rate, processing time, plate characteristics, and so forth. So this study will assist the recent researchers in the ALPR application areas to distinguish advanced techniques and their performance. This article tried to show all the state of the art techniques which is used in all the stages of ALPR.

The future research of ALPR should concentrate on multi type vehicle plate recognition, video-based ALPR, multi resolution image processing, Different character length plates etc. ALPR algorithms addressed were country specific especially for Brazilian due to availability of datasets, restricted to the working conditions such as, distance, background, illumination, vehicle position and some of the authors concentrate only on car license plates . So the future scope of research should concentrate on the prediction of LP's on the various countries and complex scene images. The researcher's should involve increasing the dataset that consist of various countries and different characteristics for vehicles including all the transportation vehicles and License plates.

REFERENCES

[1] R. Laroca, L. A. Zanlorensi, G. R. Gonc,alves, E. Todt,W. R. Schwartz, and D. Menotti, "An efficient and layout-independent automatic license plate recognition system based on the YOLO detector," arXiv preprint, vol. arXiv:1909.01754, pp. 1–14, 2019.

[2] Rabiah Al-qudah, Ching Y. Suen. **Enhancing YOLO Deep Networks for the Detection of License Plates in Complex Scenes.** International Conference on Data Science, E-learning and Information Systems 2019, ACM 978-1-4503-7284-8/19/12.

- [3] Wang, F., Zhao, L., Li, X., Wang, X., Tao, D. **Geometry-Aware Scene Text Detection with Instance Transformation Network**. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1381–1389. Salt Lake City (2018).
- [4] Xie, L., Ahmad, T., Jin, L., Liu, Y., Zhang, S. **A New CNN-Based Method for Multi-Directional Car License Plate Detection**. IEEE Transactions on Intelligent Transportation Systems 19(2), 507–517 (feb 2018).
- [5] Joseph Redmon and Ali Farhadi. **Yolov3: An incremental improvement**. CoRR, abs/1804.02767, 2018
- [6] Laroca, R., Severo, E., Zanlorensi, L.A., Oliveira, L.S., Goncalves, G.R., Schwartz, W.R., Menotti, D. **A robust real-time automatic license plate recognition based on the YOLO detector**. CoRR abs/1802.09567 (2018).
- [7] Silva S.M., Jung C.R. “**License Plate Detection and Recognition in Unconstrained Scenarios.**” In: Ferrari V., Hebert M., Sminchisescu C., Weiss Y. (eds) Computer Vision – ECCV 2018. Lecture Notes in Computer Science, vol 11216, pp 593–609. Springer, Cham, DOI: 10.1007/978-3-030-01258-8_36 (2018).
- [8] Silva, S.M., Jung, C.R.: **Real-time Brazilian license plate detection and recognition using deep convolutional neural networks**. In: 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). pp. 55–62 (Oct 2017).
- [9] Hsu, G.S., Ambikapathi, A., Chung, S.L., Su, C.P. **Robust license plate detection in the wild**. In: 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance. pp. 1–6. No., IEEE (aug 2017).
- [10] Redmon, J., Farhadi, A. **YOLO9000: Better, Faster, Stronger**. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 6517–6525. IEEE (July 2017).
- [11] Delmar Kurpiel, F., Minetto, R., Nassu, B.T. **Convolutional neural networks for license plate detection in images**. In: 2017 IEEE International Conference on Image Processing (ICIP). pp. 3395–3399. IEEE (sep 2017).
- [12] Ren, S., He, K., Girshick, R., Sun, J. **Faster R-CNN: Towards RealTime Object Detection with Region Proposal Networks**. IEEE Transactions on Pattern Analysis and Machine Intelligence , 1137–1149 (jun 2017).
- [13] Gupta, A., Vedaldi, A., Zisserman, A.: **Synthetic data for text localisation in natural images**. In: IEEE Conference on Computer Vision and Pattern Recognition (2016).
- [14] Gabriel R. Goncalves, Matheus A. Diniz, Rayson Laroca, David Menotti, William Robson Schwartz. **Real-time Automatic License Plate Recognition through Deep Multi-Task Networks**. 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, 2018, pp. 110- 117.
- [15] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. **You Only Look Once: Unified, Real-Time Object Detection**. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 779–788. IEEE (jun 2016).
- [16] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C. **SSD: Single Shot MultiBox Detector**. pp. 21–37 (2016).
- [17] He, K., Zhang, X., Ren, S., Sun, J. **Deep Residual Learning for Image Recognition**. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). vol. 4, pp. 770–778. IEEE (jun 2016).
- [18] Goncalves, G.R., da Silva, S.P.G., Menotti, D., Schwartz, W.R. **Benchmark for license plate character segmentation**. Journal of Electronic Imaging 25(5), 1–5 (2016).
- [19] Jaderberg, M., Simonyan, K., Zisserman, A., Kavukcuoglu, K. **Spatial transformer networks**. In: Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., Garnett, R. (eds.) Advances in Neural Information Processing Systems 28, pp. 2017–2025. Curran Associates, Inc. (2015).
- [20] R. Girshick, “**Fast r-cnn**,” in International Conference on Computer Vision (ICCV), 2015.
- [21] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “**Rich feature hierarchies for accurate object detection and semantic segmentation**,” in Conference on Computer Vision and Pattern Recognition (CVPR), 2014.
- [22] Jaderberg, M., Simonyan, K., Vedaldi, A., Zisserman, A. **Synthetic Data and Artificial Neural Networks for Natural Scene Text Recognition**. NIPS, Conference on Neural Information Processing Systems pp. 1–10 (2014).
- [23] Du, S., Ibrahim, M., Shehata, M., Badawy, W. **Automatic License Plate Recognition (ALPR): A State-of-the-Art Review**. IEEE Transactions on Circuits and Systems for Video Technology 23(2), 311–325 (feb 2013).
- [24] Krause, J., Stark, M., Deng, J., Fei-Fei, L. **3d object representations for finegrained categorization**. In: 4th International IEEE Workshop on 3D Representation and Recognition (3dRR-13). Sydney, Australia (2013).
- [25] Hsu, G.S., Chen, J.C., Chung, Y.Z. **Application-Oriented License Plate Recognition**. IEEE Transactions on Vehicular Technology 62(2), 552–561 (feb 2013).



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

Impact Factor:
7.512

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details