iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VII **Month of publication:** July 2022

DOI: https://doi.org/10.22214/ijraset.2022.46021

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

YOLO and Mask R-CNN for Vehicle Number Plate Identification

Siddharth Ganjoo
Department of Computer Engineering, Jamia Millia Islamia University, Delhi, India

Abstract: ***License plate scanners have grown in popularity in parking lots during the past few years. In order to quickly identify license plates, traditional plate recognition devices used in parking lots employ a fixed source of light and shooting angles. For skewed angles, such as license plate images taken with ultra-wide angle or fisheye lenses, deformation of the license plate recognition plate can also be quite severe, impairing the ability of standard license plate recognition systems to identify the plate. Mask RCNN gadget that may be utilized for oblique pictures and various shooting angles. The results of the experiments show that the suggested design will be capable of classifying license plates with bevel angles larger than 0/60. Character recognition using the suggested Mask R-CNN approach has advanced significantly as well. The proposed Mask R-CNN method has also achieved significant progress in character recognition, which is tilted more than 45 degrees as compared to the strategy of employing the YOLOv2 model. Experiment results also suggest that the methodology presented in the open data plate collecting is better than other techniques (known as the AOLP dataset).***

Keywords: Deep Learning; Mask R-CNN; YOLO; Vehicle Number Plate Identification System

I. INTRODUCTION

Because of the growth of smart towns, license plate recognition devices are planned as intelligent traffic monitoring, automatic charging, or crime detection systems that can be built into intersection sensors or street lights. License plate recognition devices in the parking lots have recently been commonly used. The conventional parking lot license plate recognition device is equipped with a fixed shooting angle and light source to quickly distinguish license plates. The standard plate recognition device first detects the license plate, then horizontally and vertically aligns, then segments the character to cut every individual character. Finally, the identification of characters takes place making use of template matching or machine learning (for example SVM and KNN) or deep learning methods (for example CNN, and DNN). On the other hand, on the intersection display, the operation of the license plate recognition device is entirely different. In the parking lot, the license plate recognition system cannot fulfill intersectional monitor criteria, such as different shot angles, blurred license plates, light and shadow shifts, etc. For instance, the proportion of the license plates recognized as well as license plate characters present in the image is very low given the high-resolution wide-angle image as seen in Fig. 1. Moreover, for tilting angles especially, as seen in Fig 2: photographs of the license plate which is taken using a super wide-angle lens or of a fisheye lens. License plate deformations can be especially severely affected, with the consequence that existing license plate identification systems are poorly recognized. Moreover, complex outdoor conditions can also impact the light source and cause excessive exposure or underexposure issues.

In the proposed paper we already have established a 3-stage R-CNN[1] license plates method, which will be useful for different shooting angles, distances as well as oblique images which helps in solving the problems mentioned above. The proposed architecture has 3 phases which include vehicle detection, license plate localization as well as character recognition. A key benefit of a 3-phase strategy is that it is not necessary to implement the phases to be used. Since the portion of the license plate of a particular image is quite closely linked to the shooting distance, one-phase architecture which isn't as helpful for training is difficult to match all samples. It is extremely hard to detect by making use of single-phase architecture considering license plates that have the longest shooting distance as well as a small portion of the frame. The three-stage architecture allows phased detection to improve the percentage of license plate characteristics in the photograph. It can be applied without taking the distance issue into account, facilitating and improving preparation. Mask R-CNN's overlapping portions of large-angular shoots are exceptional at instance segmentation. Mask R-CNN can be described as a multi-scale, pixel-pixel architecture, thus having a strong object-to-fact recognition ability.

Therefore, although the picture on the license plate is blurred, it could clearly distinguish, and correctly identify the characteristics on the license plate. The proposed architecture will describe the license plate with angles of 0~60 degrees at which the mapping rate is up to 91 percent and can be achieved with experimental performance.

©IJRASET: All Rights are Reserved

The proposed Mask R-CNN system had already included considerable improvement by determining characters that are greater than 45 degrees compared to the approach using YOLOv2[16]. The test results also demonstrate that the approach proposed is much superior to all other methods that are in the open plate dataset (known as the AOLP dataset).



Figure:1Example of the high-resolution widened-angle image

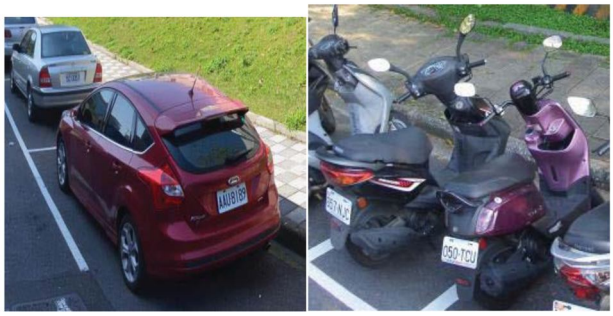


Figure:2 Deformed license plates caused due to the shooting angles

II. LITERATURE REVIEW

A. Traditional license plate recognition system

Traditional license plate recognition systems use image processing to remove appearances like horizontal and vertical projection, and edge detection from license plates.

These characteristics are recognizable primarily by naked human naked eyes and modeled mathematically on algorithms directed towards generating characteristics which humans want. Finally, character grading is performed through the contrast of features. The classification of character can be divided as follows into four methods.

1) *Template Matchin:* The matching template[8][9] is based on the distance from the regular templates calculated. Only basic screenplays including set light sources and shooting angles are required for template matches. Passing templates is fast, but changes in the shooting environment cannot be made.
2) *Machine Learning Classification:* Classifying characters directly by methods like SVM[10] as well as KNN[11]. For classification purposes, aforementioned approach use linear equation. Although the procedure is straightforward, nonlinear and multi-categories are difficult to distinguish and there is a considerable amount of estimation, that has progressively been substituted.
3) *Neural Networks:* Neural network[12][13] is programmed via derived characteristics, as conventional approaches will, so computer can mimic human beings in order to understand these characteristics.
4) *Convolutional Neural Networks:* Conventional plate recognition devicethat can extract human characteristics which is used for training machines are not as simple to learn and classify, even though it appears intuitive to humans, and their effects are no better than the machine itself. The proposal for the formation of a system to remove functionality and identify objects by itself[14][14] is then made for coevolutionary neural networks(CNN). While CNN can isolate the features itself, only photographs can be classified and the effect of object location cannot be achieved. It is only possible to identify one character each on a license plate in the final phase of the standard license plate recognition scheme.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

B. Object detection-based license plate recognition system

A license plate recognition system based on object recognition [16][17] is used for locating and detecting characters present on license plates using object detection master-learning architecture. While the license plate identification systems focused on object detection had significantly increased noise capability as well as broad-angle recognition capability compared with conventional license plate recognition systems, the segmentation capability of overlapping characters remains inadequate.

1) *YOLOv2:* YOLO[7] also known as You Look Once, implies that for the whole picture, a one-story architecture, only one shot is detected. YOLOv2 has several anchor boxes of various sizes, with each picture split into s*s. Any center of the grid is handled one at a time as an anchor, and all anchor boxes are then detected. YOLOv2 improves the generalization through batch normalization and increases YOLO's mAP rate with several anchor boxes, high-resolution classifiers, a cluster of dimensions, a precise position predictor, training, and sophisticated features. YOLOv2 is generally fasters and more precise than YOLO, thus improving the precise identification of small objects.
2) *Mask R-CNN:* The mixture of ResNet[2] as well as FPN[4] could be contemplated loosely as the mask of R-CNN. The Faster R-CN N[3] as well as FCN[5] combination and part of the mask formed by FCN can be considered.

a) *ResNet:* ResNet's key benefit is the degradation solution to Deeper neural teaching network issues. ResNet is a multi-layered architecture consisting principally of four remaining blocks which match each other repeatedly Layers. - Layers. The remaining block creates an entry shortcut with two coevolutionary layers each with a sophisticated feature added to the outputs to maintain the power of initial entry and resolve the deep plain issue It is impossible to stably converge networks.

b) *Faster R-CNN:* Faster R-CNN is known as2-stage method, with a phase of proposal and phase of identification. The backbone of the Faster R-CNN is ResNet and the projected bounding boxes can be found with RPN. Faster R-CNN produces feature maps by ResNet in the proposal phase and that dispatches feature maps for Region Proposal Network (RPN). RPN is a comprehensive network that uses a sliding window for defining zone proposals by using anchor boxes of varying sizes. Regions ideas are returned to the original ResNet performance charts in the detection stage, RoIPool is performed, the RoI is corrected, the final bounding box is retrieved, and classes, as well as boxes, are output.

c) *Feature Pyramid Network (FPN):* For effects of convolution and pooling layers, FPN produces various scales of function maps. The bottom feature map makes the RPN performance forecast, samples up and adds to the previous feature map 4 times bigger than before, and generates the RPN projection output again. The same process is repeated by adding an element to the previous 4-fold function map and predicting the RPN performance then.

d) *Fully Convolutional Networks (FCN):* FCN is a completely coevolutionary network. Unlike conventional CNN, FCN is not completely bound to layer. Each pixel in the image is explicitly labeled and classified by FCN such that semantic segmentation is possible.

e) *RoIAlign:* To optimize Faster R-CNN, RoIAlign is used. RCNN accomplishes pooled and quantized ROI quicker. Since the RoI area does not land on the grid junction, spatial offsets are caused. This may not be a significant issue at first sight, but the precision of detection of small items has been significantly affected. The major contribution of Mask R-CNN is to resolve the problem which is caused due to pooled operations performed on RoI RoIAlign is used to replace RoIPool with Mask R-CNN. RoIAlignwill not quantify RoI explicitly, but raises the sampling point of each container and by bilinear interpolation measures, the value of each sampling point and then extracts a total RoI value. Thus, RoIAlign prevents offsets and errors which are caused by quantization.This is the R-CNN mask definition. Since Mask R-CNN employs many different network optimization strategies, which has many benefits. E.g., ResNetstabilises the R-CNN mask for best training performance.RPN can track objects and locate boxes in Faster R-CNN. To achieve instance segmentation, FCN will perform pixel-pixel training along with the correction. R-CNN mask attaches FCN to the backside of RPN moreover only executes FCN present on side of its bordering box for secure instance.RoIAlign is more exact than other methods of RoI of Mask R- CNN. The FPN allows the detection effect of small objects strong in multi-scale, significantly better than the detection effect.

III. Methodology

In this segment, we suggest a 3-phase license plate recognition scheme involving vehicle detection then license plate localization furthermore character recognition, as seen in Fig. 3.

A. Vehicle Detection

At the initial stage, YOLOv2 is adopted for detecting vehicles. Vehicles listed include cars, lorries, buses, and trucks to meet real needs. The picture is split into a grid of 19*19.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Since the car is of various proportions, the detection frame is made up of five types of anchor boxes. This phase is used to determine the vehicles to be included in a picture and then to mark the areas of the identified vehicles under the name RoI (area of interest).

B. License Plate Localization

We also use YOLOv2 for license plate detection in the second round. In stage I, YOLOv2 division into the first 19 x 19 grid. Because of the skewed angle of the license, five types of anchor boxes are applied, these anchor boxes are used for every license plate detection center, and afterwards, for character identification, the license plate images are captured in the final process.

C. Character Recognition

In the last stage, we take benefit of the R-CNN mask to recognize character. On the grounds that the R-CNN mask is architecture basically useful for each segmentation, the edges of each character must be labeled so that they can detect pixels by pixels. We have chosen ResNet-50 known as the backbone of Mask R-CNN in ResNetwhich is based on a trade-off between speed and precision. The entire convolution network is useful for RoI masking after object recognition is done. Finally, the character on the license plates that were caught in the last phase is identified using Mask RCNN.

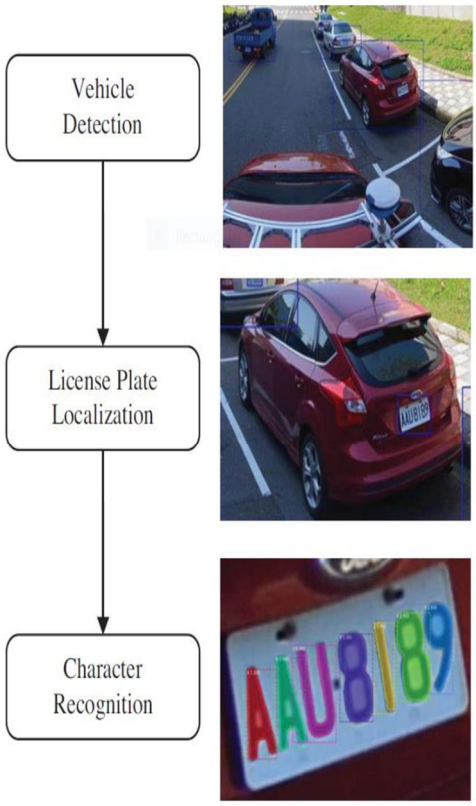


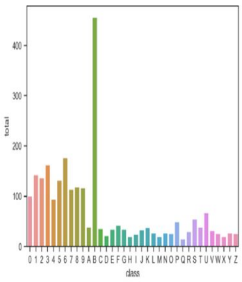
Figure:3Stages of proposed license plate recognition system

Generally speaking, proposedlicense plate identification device should no more have to be used together in three stages and can be efficiently deployed. For instance, Phase III could be useful in directly to recognize character even without locating the vehicle detected moreoverlicense plate if only license plate occupies more than 20% of the picture. If the license plate ratio for photographs range from 20% to 3%, PhaseII for the identification of license plates and PhaseIII for classification of characters without vehicles will be used. Entire three-phase scheme will be implemented if only ratio of license plate to photographs is lesser than 3%. It should be noted that if ratio of license platform to picture is poor, it can be understood as simple to miss license platesalong with characters which contains one-stage profound learning architecture. In the phases I and II, YOLOv2 is used to collect vehicles and license plates, so that the character/image ratio is increased. Thus, in third step we will increase the reminder rate of the identification of the character and reduce significantly the problem that license plate characters cannot be detected.

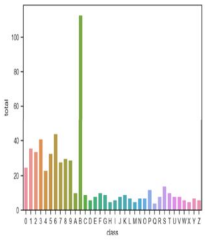
©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV. EXPERIMENTS

Fig. 4 shows extent of preparation along withapproval information required for ResNet model utilized in detection of License Plate digit classification.



(a) Training



(b) Validation

Figure:4 Total amount of data perc class for training as well as validation of ResNet Model.

The first step is to take the YOLOv4 Bounding Coordinates and simply take the region of the sub- image within the boundaries of the box. Since we utilize cv2.resize() for blowing the picture up to 3x its original size, most of the time it is incredibly tiny



We next transform the picture in grey and smooth it by applying a slight Gaussian blur.



After this, the image has black background threshold to white text and the Otsu technique has been applied. This black white writing helps to detect image contours.



The picture is then dilated using RCNN mask to increase the visibility of contours and to be taken in the future.



Then we use Mask-RCNN to detect all of the rectangular contours on the picture and sort them from left to right.
As we can see, several contours are contained inside the license plate number other than the contours of each character. We use a handful of characteristics to accept a contour in order to filter out the undesirable parts. These are height and width (i.e. the area height must be 1/6th the whole picture height) alone. These are height and width ratios. A few other characteristics are also set on the area etc. To view full details, check out the code. We are left with this filtering.



The only areas of interest remaining now are the individual characters of the license plate number. We segment and apply a bitwise not mask on the white backdrop of every picture in order to flip the picture into black text with which Tesseract is precise. A slight median blur on the image is the final stage, then the letter or number from it is given on to Tesseract. Example how letters would look when applied to tesseract.



Each letter or number will then only be joined together in a string and you will finally receive the entire recognized license plate. The experimental equipment's usedare CPU of Intel Core i7-4790 alongside GPU of NVIDIA, GeForce GTX TITAN X. Intel Core i7-4790 which contains of cores operate at 3.60GHz while TITAN X which has 3,840 CUDA cores which operate at 1GHz. Operating system used is Linux Ubuntu 16.04 LTS. Test item is general car license plate. The number plate consists of a mixture of A~Z (except I, O) and 0~9 for a total of 34 characters. In addition, the total number plate is 8. There are a minimum of 4 and a limit of 7 license plates on each license platform. The license plate's background colour is white, black, red or green. However number of currently issued license plates vary somewhat based on the number of cars, the bulk of these are black and white license plates. In the sunny and snowy days, the pictures were taken. There are two sections of the experimental findings. The pictures taken from roadside along with road bridge are usefulin training YOLOv2 model for vehicle identification and license plate detection. Multiple trails and multiple vehicles can include a picture. The size of the input image is 1920 pixels. The training data set which consists of 4,500 images and validation data set which consists of 500 images. These pictures include various vertical and horizontal offset and rotated angle. For license plate with angle range above 0~60 degrees, the characters recognition mAPconsidered for validation dataset is about 91%.

©IJRASET: All Rights are Reserved



Figure:4 License plate examples in training and testing dataset



Figure:5License plate patternswhich were taken from various angles

V. RESULTS

Figure:6 shows illustrations of misfortune as well as exactness of model. During preparation, we save loads which returns with best approval exactness as well as utilization for tag acknowledgment framework. The model has 0.80159 approval exactness.

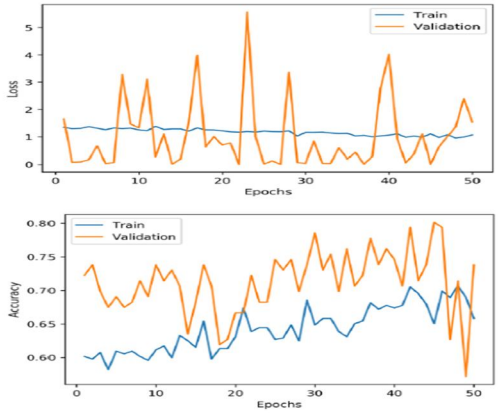


Figure:6 The graphics for loss alongside accuracy of ResNet model

In Table I, Recall rate for each and every angular range begins at the normal position of lens aligned to normal direction of centre plate, each of which is 15 degrees, according to figure 5. The reminder rate of each angled range. The study contains a left-hand side path and a right-hand path, each 50 percent. As the photographs are captured from real use, they can be concentrated at a distance of 150 ~ 200 cm, so the detection effect of 50cm is comparatively low. According to the real use, our data collection also contains a limited amount for samples from 60 to 75 degrees.The time of processing of each and every image is the result of image recognition. Detection time would be abnormally quick if image recognition effect is not successful. In Table I, for example, images of 45 to 60 and 60 to 75 degrees are processed in short periods than the images of less than 45 degree images. Processing time of 50 cm images is therefore less than the processed time of 100-200 cm images. In Table I, the proposed approach with RCNN mask shows major advances when it comes to characters inclined above 45 degrees compared to method using the YOLOv2 model. For example proposed approach of Mask R-CNN achieves 76.8 per cent accuracy for photographs that are tilted from 45 directed towards 60 degrees over distance of about 200 cm, whereas YOLOv2 method is just 48 percent accurate.

Model	A two-stage approach [16] using YOLOv2: License Plate Localization + Character Recognition				The proposed approach with Mask R-CNN				Average Improvement
Angle \ Distance	50cm	100cm	150cm	200cm	50cm	100cm	150cm	200cm	
0 ~15	8.6%	83.9%	76.2%	99.9%	87.1%	97.9%	99.9%	99.9%	29%
15 ~30	17.0%	90.7%	98.7%	99.8%	80.5%	94.2%	97.2%	99.9%	16%
30 ~45	0.4%	85.7%	84.1%	93.1%	63.7%	85.6%	93.6%	98.6%	19%
45 ~60	0.1%	48.2%	37.6%	48.0%	25.5%	62.1%	79.6%	76.8%	28%
60 ~75	0.0%	0.1%	4.0%	4.2%	6.4%	13.1%	26.7%	22.1%	15%

Table I: Comparison of recall rate of each and everyangular range.

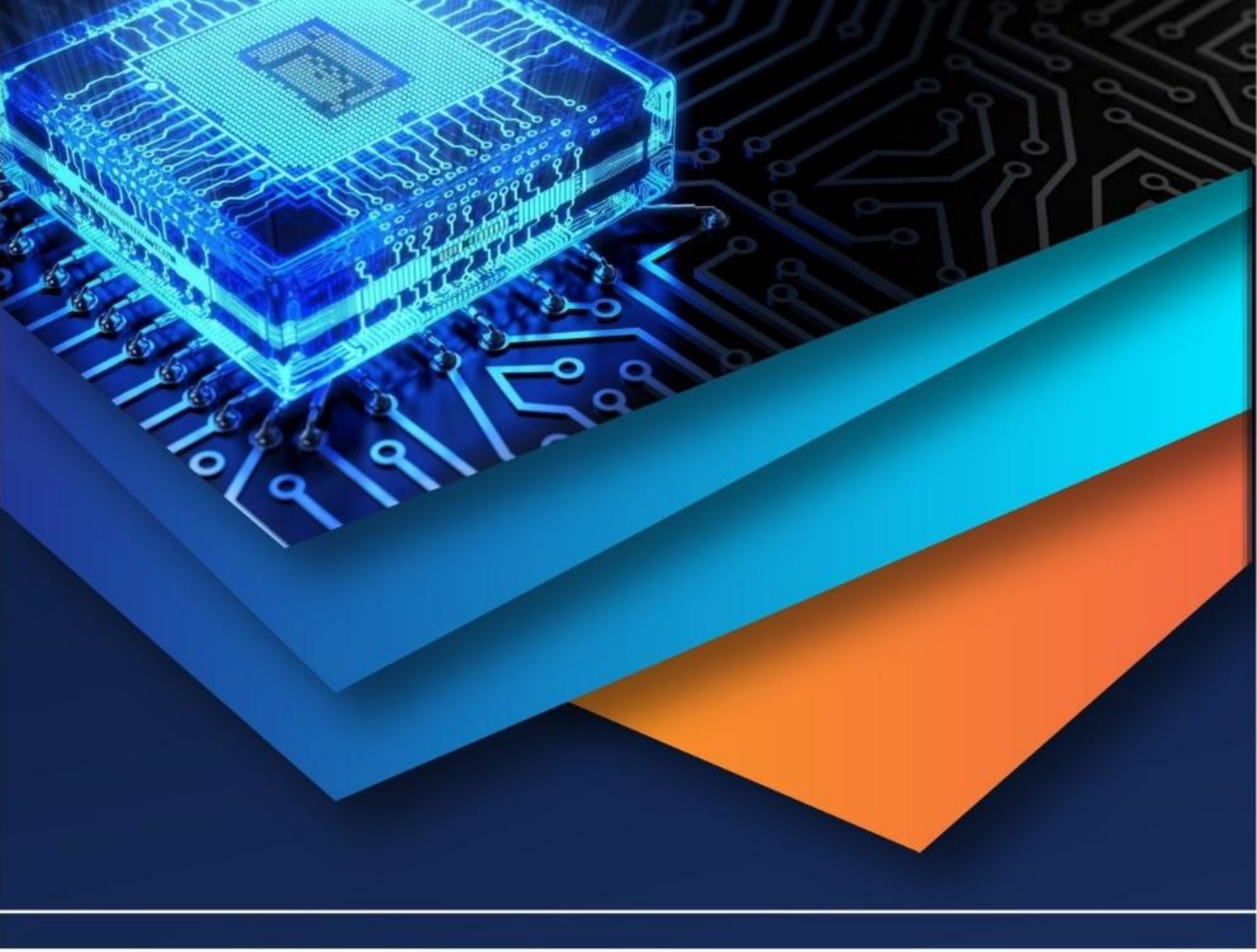
VI. CONCLUSION

In the proposed study, we present a three-stage R-CNN mask-based licence platform identification approach that may be applicable to higher oblique and diverse filming angles. According on experimental results, the proposed design will differentiate licence plates with radius angles between 0 and 60. When comparing the suggested R-CNN mask technique to the YOLOv2 model, significant progress has been made in the detection of characters that are slanted at an angle greater than 45 degrees.

REFERENCES

[1] K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 2980-2988.
[2] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 770-778.
[3] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017.
[4] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 936-944.
[5] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 3431-3440.
[6] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 6517-6525.J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once:Unified, Real-Time Object Detection", CVPR, LAS VEGAS, 2016, pp.1-10.
[7] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 779-788.
[8] M. Yu and Y. Kim. "An Approach to Korean License Plate Recognition Based on Vertical Edge Matching," SMC, 2000 , pp.1-6.
[9] J. Wang, W. Zhou, J. Xue and X. Liu, "The research and realization of vehicle license plate character segmentation and recognition technology," 2010 International Conference on Wavelet Analysis and Pattern Recognition, Qingdao, 2010, pp. 101-104.
[10] Y. Wen, Y. Lu, J. Yan, Z. Zhou, K. M. von Deneen and P. Shi, "An Algorithm for License Plate Recognition Applied to Intelligent Transportation System," IEEE Transactions on Intelligent Transportation Systems (T-ITS), vol. 12, no. 3, pp. 830-845, Sept. 2011.
[11] S. Karasu, A. Altan, Z. Sarac,and R. Hacioglu,"Histogram based vehicle license plate recognition with KNN method," ICAT, Rifat, 2017, pp. 1-4.
[12] C. N. E. Anagnostopoulos, I. E. Anagnostopoulos, V. Loumos and E. Kayafas, "A License Plate-Recognition Algorithm for Intelligent Transportation System Applications," IEEE Transactions on Intelligent Transportation Systems (T-ITS), vol. 7, no. 3, pp. 377-392, Sept. 2006.
[13] T. Nukano, M. Fukumi, and M. Khalid, "Vehicle license plate character recognition by neural networks," Proceedings of 2004 International Symposium on Intelligent Signal Processing and Communication Systems 2004 (ISPACS 2004), Seoul, South Korea, 2004, pp. 771-775.
[14] S. Li and Y. Li, "A Recognition Algorithm for Similar Characters on License Plates Based on Improved CNN," 2015 11th International Conference on Computational Intelligence and Security (CIS), Shenzhen, 2015, pp. 1-4.
[15] 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.
[16] C. Lin and Y. Sie, "Two-Stage License Plate Recognition System Using Deep learning, " ICASI, 2019, pp. 1-4.
[17] S. Abdullah, M. Mahedi Hasan and S. Muhammad Saiful Islam, "YOLO-Based Three-Stage Network for Bangla License Plate Recognition in Dhaka Metropolitan City," 2018 International Conference on Bangla Speech and Language Processing (ICBSLP), Sylhet, 2018, pp. 1-6.
[18] G. Hsu, J. Chen and Y. Chung, "Application-Oriented License Plate Recognition," in IEEE Transactions on Vehicular Technology, vol. 62, no. 2, pp. 552-561, Feb. 2013.
[19] J. Jiao, Q. Ye, and Q. Huang, "A configurable method for multi style license plate recognition, "Pattern Recognit.,2009, pp.358–369.
[20] C. N. E. Anagnostopoulos, I. E. Anagnostopoulos, V. Loumos and E. Kayafas, "A License Plate-Recognition Algorithm for Intelligent Transportation System Applications," in IEEE Transactions on Intelligent Transportation Systems, vol. 7, no. 3, pp. 377-392, Sept. 2006.
[21] Z. Selmi, M. Ben Halima and A. M. Alimi, "Deep Learning System for Automatic License Plate Detection and Recognition," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, 2017, pp. 1132-1138.ssss

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)