2023 Global Conference on Information Technologies and Communications (GCITC)
Karnataka, India. Dec 1-3, 2023

Real-Time License Plate Detection and Recognition System using YOLOv7x and EasyOCR

Sachin Dhyani
Robotics and Mechatronics Research Lab
Chitkara University Institute of Engineering and Technology,
Chitkara University,
Punjab, India,140417
sachin.2387@chitkara.edu.in

Vijay Kumar
Robotics and Mechatronics Research Lab
Chitkara University Institute of Engineering and Technology,
Chitkara University,
Punjab, India,140417
vijay.jadon@chitkara.edu.in

***Abstract*— Automated license plate detection system is used for recognition and locating of license plates on vehicles using image processing and computer vision methods. Automated license plate detection is an important area of research in computer vision with far- reaching applications in the industrial sector. It has the potential to increase productivity and accuracy in numerous industrial applications. The procedure entails utilizing a camera to take a photo of a vehicle's license plate, processing the photo to identify the plate, and extracting the characters from the plate. This paper provides implementation of the YOLOv7x algorithm to detect license plate on vehicles. EasyOCR is applied to the detected license plate to recognize the text. The model is trained on a customized dataset consisting of only Indian vehicles license plate. The proposed work achieves a high accuracy of 99% and outperforms old techniques like Yolov3, darknet, CNN.**

Keywords*— *autonomous vehicle, license plate recognition, object detection, yolov7.

I. INTRODUCTION

Road safety is one of the most important issues in the advancement of smart cities when it concerns to autonomous mobility[1]. Every vehicle on road is affixed with a license plate on its front & back side[2]. The license plate helps in identification purposes as it contains the information about the owner of vehicle[3]. There are many potential applications for vehicle identification and recognition in automatic driving, making it a hot research area in computer vision[4]. Automatic license plate recognition system can be helpful in various scenarios such as border security, traffic monitoring, electronic toll collection, theft vehicle identification and parking lots[5].

The automatic license plate recognition system can be categorized into three segments: detecting the license plate, segmentation of license plate and recognizing the characters from the license plate[6]. For the detection of license plates various object detection algorithms are used such as CNN, SSD, Faster R-CNN, Yolo[19]. The detected license plate can be recognized using OCR, character segmentation and training a deep learning model for character recognition. License plate detection and recognition can be challenging due to the various factors.

License plate detection and recognition can be challenging due to the various factors. The factors that must be taken into consideration are specific detectors or viewing angles, appropriate lighting, different types of vehicles and variety of license plate font[7]. The position of camera and distance from the target impacts the process of real-time image collection and results in issues including block, blur, dim light, and tiny size of the target item[8]. A generalized character recognition system may not always function well as some characters can be misidentified and result in low recognition rate. However, the advancements in deep learning have changed the approach used to extract picture features[20]. M.A.R. Refat et al. presented a 16-layer effective CNN method for classifying human face expressions using data augmentation. Also, the suggested method surpasses several earlier studies with state- of-the-art testing accuracy of 89.89%[9]. The methods based on deep learning can extract features by themselves and learn, in contrast to typical manual feature extraction[10].

YOLO (You Only Look Once) algorithm was first introduced to the world by Redmon et al.[11] in 2016. Due to its superior performance over other object detection techniques such as CNN, the YOLO algorithm has grown in popularity. Jing Han et al. [12] suggested a CNN-based multi- oriented and scale invariant license plate detection model that can outperform current methods in terms of detecting license plates by drawing bounding parallelogram rather than conventional horizontal rectangular boxes over region of interest. N Sharma et al.[13] proposed a support vector machine and CNN-based license plate recognition system in which they retrieved characters and segmented them from input image of car and extract the features from it. The extracted information is passed through CNN and support vector machines to categorize the extracted information for the final recognition of the license plate. V Gnanaprakash et al.[14] implemented YOLO for license plate detection by identifying the car from the video frames in the first stage. In the second stage License plates are detected from the images and the number plate characters are recognized from the observed number plate. Their deep learning model makes use of the Image AI library to facilitate training. S Agrawal et al.[15] used YOLOv7 and scene text recognition for detection and recognition of license plates in automatic weighbridge services. For many image-based detection strategies, the segmentation techniques are also used but image segmentation is widely employed in the identification and treatment of human diseases[16].

Hong-yuan Mark Liao et al.[17] were the first to introduce YOLOv7 to the world. A generalized network architecture of the model can be seen in Fig.1. A fresh approach for object detection in pictures or videos is the YOLOv7x algorithm. In comparison to previous neural networks, YOLOv7x can be trained faster on smaller amounts of data without relying on prior knowledge. The YOLOv7 model consists of a backbone and head whereas the previous versions also had a neck. The input is first provided to the backbone. The backbone is responsible for extracting features from the input image. The backbone network uses a succession of convolutional procedures to extract high-level features at various levels of abstraction from the input image.

It consists of numerous sequentially connected convolutional layers used to apply convolutional operations. Further it also consists of maxpooling layer that downscale the input image's spatial dimensions while retaining important features and concatenation layers that combines the outputs from multiple preceding layers. The backbone produces a collection of feature maps containing both local and global data and depict the image at various sizes. These feature maps are then provided as input to the head network. It first applies Spatial pyramid pooling to capture multi-scale information. Convolutional, upsampling, and concatenation layers are among the layers that make up the head. To provide the final estimates, it merges the feature maps from various resolutions. A series of bounding box predictions, along with the corresponding class labels and confidence scores, are the output of the head. These predictions are produced using feature maps and predetermined anchor boxes.

In our proposed method, a YOLOv7x model is trained to detect license plate and then extract the text utilizing EasyOCR. The rest of the article is assembled as follows: the section II of our paper outlines our methodology, followed by section III containing our results and section IV serves as the conclusion of paper.

II. METHODOLOGY

Training and testing are the main procedures in this study, as indicated in Fig.2. The methodology section also describes preparation of dataset for training and validation. It also involves the extraction of frames, creating annotations of our dataset and dividing our dataset into training and validation. Furthermore, the YOLOv7x algorithm for detecting license plates is described, followed by character recognition using EasyOCR.

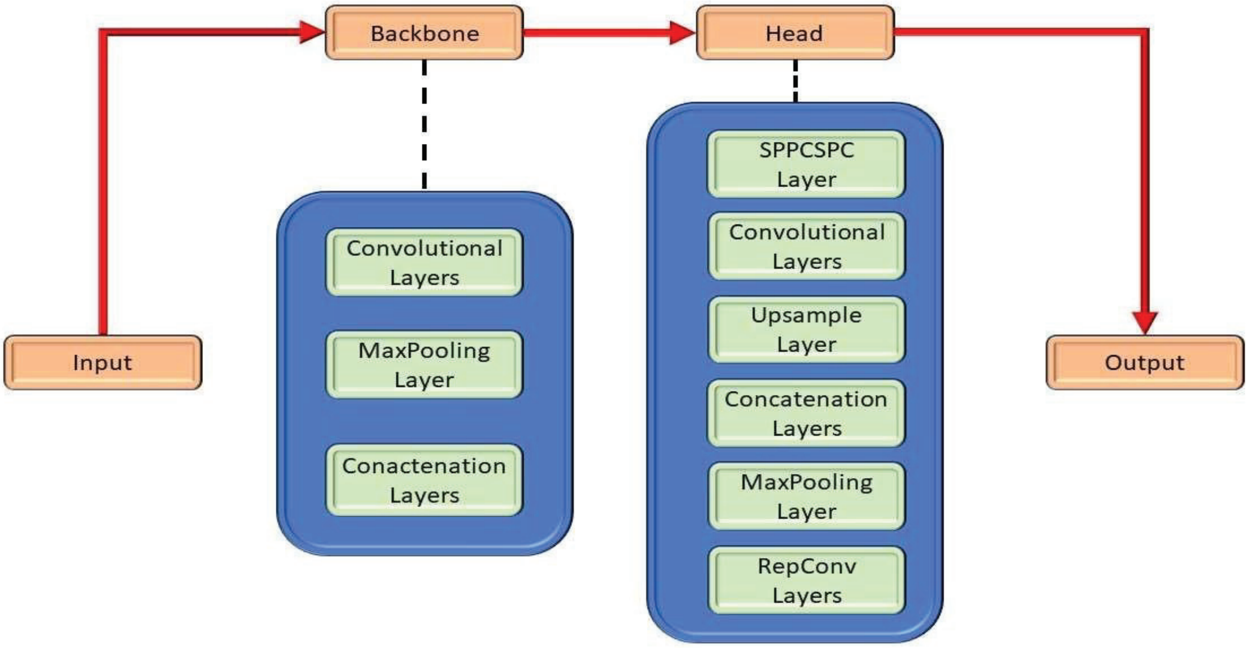

Fig.1. Model Architecture

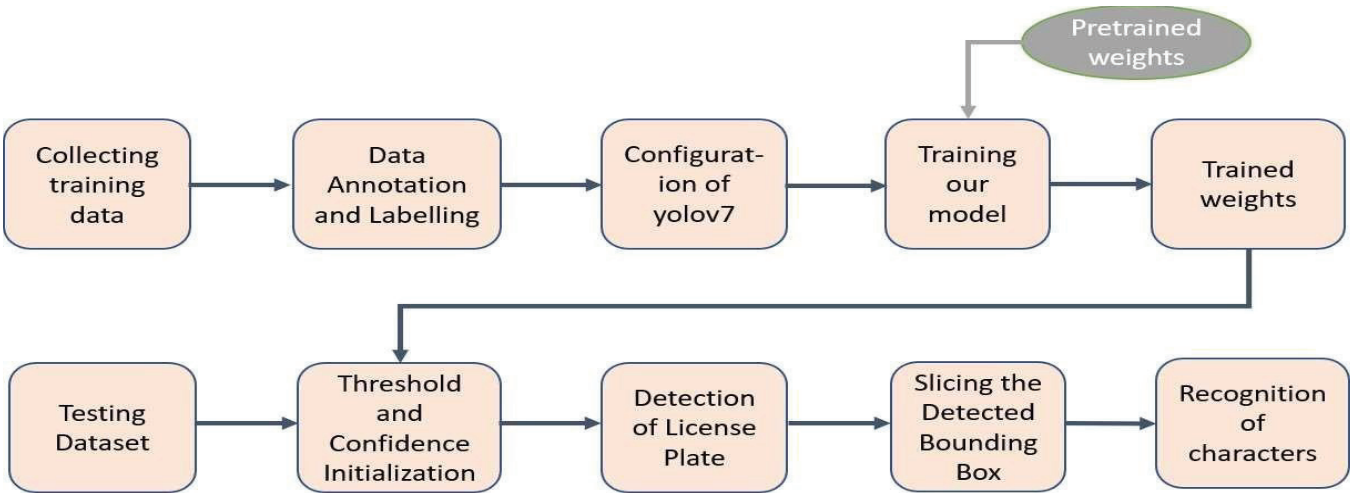

Fig. 2. Proposed method

A. *Dataset Preparation*

Training data is necessary to create weights that will be used during object recognition based on digital images. Furthermore, some data can be used for the testing procedure to check the method's accuracy or real time input can be provided to the model. For the proposed model, a custom dataset was created with pictures of vehicles self-photographed in a parking lot and other downloaded from web. The dataset consists of license plate images of only Indian vehicles. The next step is to split data between training and validation in 0.8:0.2. The dataset must be in YOLOv7x format, which means that each picture in the dataset must have a corresponding annotation. Data preprocessing and data annotation are two very important steps that needs to be done before training on the dataset.

B. *Data Preprocessing*

The image quality plays a vital role in detection of objects. Some preprocessing is required before the images are annotated. The purpose behind preprocessing the images is to improve their quality as some of them have different size, orientation, and lightning condition. The steps that are taken during preprocessing are adjusting color intensity, resizing of image, noise reduction and increasing the contrast along the edges in order to enhance the visual detail.

C. *Data Annotation*

A rectangular box is drawn over the license plate in order to annotate it. The annotation details are stored in a file. The annotation file for each image is different and contains specific details of the image such as class id, x and y coordinates of bounding box's centre, width and height of bounding box and image. The data annotation illustration is shown in Fig.3. A green rectangular box is created over the license plate.

Fig.3. Sample Image Showing Annotation of the Data

D. *License Plate Detection Using YOLOv7x*

The YOLOv7x model, which was initially pretrained on the COCO dataset, was modified for training on a custom dataset to meet certain research goals. The configuration and YAML files for the model have undergone several significant changes. These modifications included setting new values for variables like the class count and class names as well as defining the locations of the training and validation datasets. We chose to carry out the model training on Google Collaboratory, which offers free access to potent GPUs, in order to speed up the procedure and take use of GPU acceleration. Some adjustments were made to the model's hyperparameters to attain peak performance during the training phase. A batch size of 16, an image size of 640x640 pixels, and a training period lasting 150 epochs were important hyperparameters assigned. This training effort resulted in a weight file that contains the model's newfound information. This weight file serves as evidence of YOLOv7x's suitability for our dataset and study goals. For the next stage of inference, this weight file is essential since it contains specialized knowledge about our dataset. The inference results using this weight file can be seen in Fig.4. in results section.

E. *Text Recognition using EasyOCR*

EasyOCR is a powerful tool used for the purpose of reading and extracting text. It supports various languages and recognizes the text in different styles and fonts. The license plate after successful detection is cropped from the image and provided as input to the OCR. The input image to the OCR can be seen in Fig.5.

Fig.5. Input image to OCR

OCR reads the bounding box coordinates from top left corner to bottom right corner and recognize all the numbers and characters. To improve the quality of OCR results, text containing the pattern "IND" was excluded from the final annotations. This filtering step ensured that only relevant text was displayed on the annotated images. The OCR results are visualized on the images using OpenCV. YOLO can also be used for recognition of the characters but its output is not sequential i.e., it displays character whichever is recognized first[18]. To overcome this EasyOCR is used that gives the output in sequential order that matches the license plate. The output of OCR is shown in Fig.6.

III. Results And Discussion

The result obtained from the inference can be seen in Fig.4. The numberplate has been successfully detected and can be seen with a red color bounding box drawn over the license plate.

Fig.4. License Plate Detection

The text obtained from the OCR on the input bounding box image are in Fig.6.

Fig.6. License Plate Recognition

The findings of the text recognition were precise and matched the text on the license plate. Bounding boxes and text labels were added to highlight the recognized text, providing a clear representation of the OCR output.

Throughout the training, it was vital to often assess the precision, recall, and mAP parameters. After successful training of our model, we are provided with some metrices. With the help of these metrices we can analyze our model results. Plotting the precision and recall values for various classification thresholds results in the precision and recall curves. Precision gives us the count of the number of true positive cases among the anticipated positive ones. Recall, also known as sensitivity, measures the proportion of positive cases that the model genuinely recognizes. The precision-recall curve for our model is shown in Fig.7.

A precision and recall trade-off led to a high mAP score of 0.99 at a threshold value of 0.5. Therefore, the model performs well for the input class and can confidently and reliably detect item. It is crucial to keep in mind that assessing a model's performance using a single metric, such as mAP, can occasionally be inadequate because it might not capture all aspects of a model's performance. As a result, a more thorough assessment of the model's performance is required, which includes looking at other metrices as well such as confidence matrix, F1 curve. The confidence matrix is not considered as an appropriate parameter where only a single class is to be detected. In such cases, the F1 curve is analyzed.

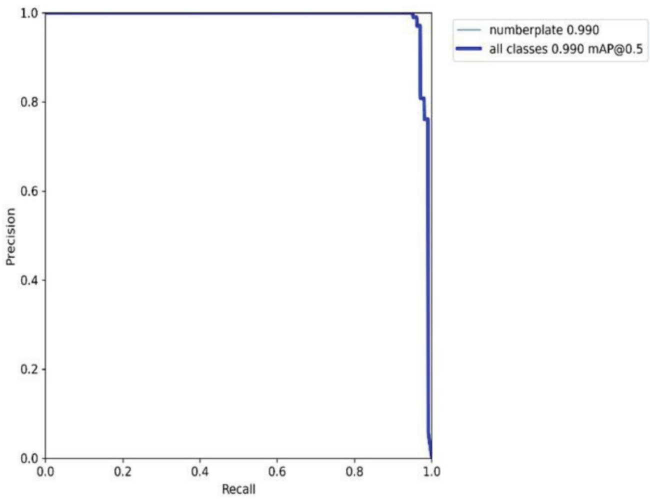

Fig.7. Precision vs Recall values plot

The F1 curve provides F1 score, a parameter used to evaluate the effectiveness of a classification model. The F1 score is a harmonic mean of recall and precision. The F1 score varies between 0 & 1. The recall and precision criteria are balanced by the F1 score, which weighs precision and recall equally. The F1 curve plots F1 score against varying confidence threshold values. If the predicted probability is greater than confidence value then it is considered as positive otherwise negative. The trade-off between precision and recall can be adjusted by varying confidence value. The F1 curve of the model is shown in Fig.8. The confidence value that maximizes precision and recall according to the F1 curve is 0.612. The F1 score obtained is 0.98, and the classification cut- off is 0.612. Witha balanced trade-off between precision and recall, our model has a good overall performance according to an F1 score of 0.98. In multi-class classification, the threshold might change depending on the problem and the model, however in binary classification, a threshold of 0.5 is frequently employed. For single class classification the threshold value is not that important but correct detection is more beneficial.

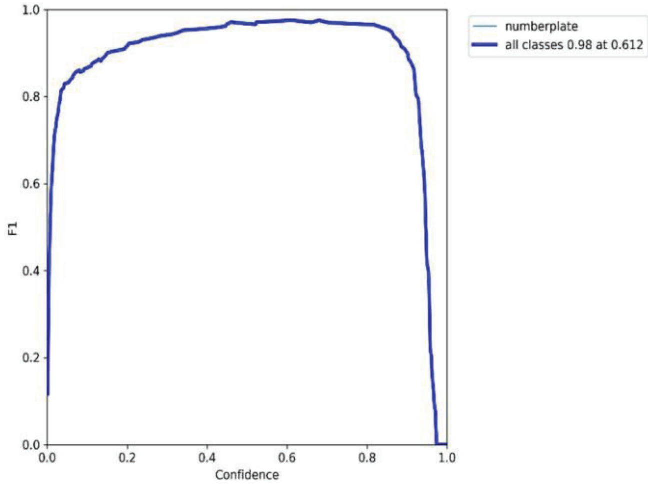

Fig.8. F1 curve of model

The metric results are summarized in TABLE I.

TABLE I. TRAINING RESULTS OF MODEL

Class	Metric	Result
License Plate	PR Curve	0.99
License Plate	F1 Curve	0.98

The comparison of our proposed method with some recent work is shown in TABLE II.

TABLE II. COMPARISON TABLE

Ref. no.	Technique	mAP
[13]	CNN & Vector-machine	96.5%
[18]	YOLOv3 & Darknet-53	80%
Proposed Work	YOLOv7x & EasyOCR	99%

IV. CONCLUSION

In this study, we have shown how the YOLOv7x algorithm can be used to detect license plate of the vehicles and then extract characters using EasyOCR. The custom dataset was used to train the neural network in order to detect the license plate on vehicle and later recognize the text on license plate. The model achieved high accuracy and outperformed older techniques. As, the solution provides accurate and reliable response it is advantageous to various groups of people and organizations such as tolling agencies, parking management companies, gas stations, etc. By including further changes like training on diverse dataset or by investigating the usage of different deep learning methods, the suggested work can be further enhanced.

REFERENCES

[1] M. Saleem, S. Abbas, T. M. Ghazal, M. A. Khan, N. Sahawneh, and M. Ahmad, "Smart cities: Fusion-based intelligent traffic congestion control system for vehicular networks using machine learning techniques," Egyptian Informatics Journal, vol. 23, no. 3, pp. 417-426, 2022.

[2] X. Zhao, Z. Huang, and Y. Lv, "Research on Real- Time Diver Detection and Tracking Method Based on YOLOv5 and DeepSORT," in 2022 IEEE International Conference on Mechatronics and Automation (ICMA), 2022: IEEE, pp. 191- 196.

[3] A. Boby and D. Brown, "Improving Licence Plate Detection Using Generative Adversarial Networks," in Pattern Recognition and Image Analysis: 10th Iberian Conference, IbPRIA 2022, Aveiro, Portugal, May 4–6, 2022, Proceedings, 2022: Springer, pp. 588-601.

[4] C.-N. E. Anagnostopoulos, I. E. Anagnostopoulos, D. Psoroulas, V. Loumos, and E. Kayafas, "License plate recognition from still images and video sequences: A survey," IEEE Transactions on intelligent transportation systems, vol. 9, no. 3, pp. 377-391, 2008.

[5] W. Riaz, A. Azeem, G. Chenqiang, Z. Yuxi, and W. Khalid, "YOLO based recognition method for automatic license plate recognition," in 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA),2020: IEEE, pp. 87-90.

[6] Z. Chen, L. Cao, and Q. Wang, "Yolov5-basedvehicle detection method for high-resolution UAV images," Mobile Information Systems, vol. 2022, 2022.

[7] Y. Zou et al., "License plate detection and recognition based on YOLOv3 and ILPRNET," Signal, Image and Video Processing, vol. 16, no. 2,pp. 473-480, 2022.

[8] T.-G. Kim et al., "Recognition of vehicle license plates based on image processing," Applied Sciences, vol. 11, no. 14, p. 6292, 2021.

[9] M. A. R. Refat, S. Sarker, C. Kaushal, A. Kaur, and M. K. Islam, "WhyMyFace: A Novel Approach to Recognize Facial Expressions Using CNN and Data Augmentations," in Emerging Technologies in Data Mining and Information Security:Proceedings of IEMIS 2022, Volume 3: Springer, 2022, pp. 553-563.

[10] 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.

[11] S. M. Silva and C. R. Jung, "Real-time licenseplate detection and recognition using deep convolutional neural networks," Journal of Visual Communication and Image Representation, vol. 71,p. 102773, 2020.

[12] J. Han, J. Yao, J. Zhao, J. Tu, and Y. Liu, "Multi- oriented and scale-invariant license plate detection based on convolutional neural networks," Sensors, vol. 19, no. 5, p. 1175, 2019.

[13] N. Sharma, P. K. Dahiya, and B. R. Marwah, "A hybrid approach for automatic licence plate recognition system," International Journal of Sensors Wireless Communications and Control, vol. 11, no. 1, pp. 66-71, 2021.

[14] V. Gnanaprakash, N. Kanthimathi, and N. Saranya, "Automatic number plate recognition using deep learning," in IOP Conference Series: Materials Science and Engineering, 2021, vol. 1084, no. 1: IOP Publishing, p. 012027.

[15] S. Agrawal and K. D. Joshi, "Indian Commercial Truck License Plate Detection and Recognition for Weighbridge Automation," arXiv preprintarXiv:2211.13194, 2022.

[16] V. Kukreja and P. Dhiman, "A Deep Neural Network based disease detection scheme for Citrus fruits," in 2020 International conference on smart electronics and communication (ICOSEC), 2020: IEEE, pp. 97-101.

[17] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv preprint arXiv:2207.02696, 2022.

[18] B. Setiyono, D. A. Amini, and D. R. Sulistyaningrum, "Number plate recognition on vehicle using YOLO-Darknet," in Journal of Physics: Conference Series, 2021, vol. 1821, no. 1: IOP Publishing, p. 012049.

[19] S. Dhyani and V. Kumar, "Multi-class Traffic Sign Recognition System Using One-Stage Detector YOLOv5s," in 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN), 2023: IEEE, pp. 1-5.

[20] U. K. Lilhore et al., "Hybrid model for detection of cervical cancer using causal analysis and machine learning techniques," Computational and Mathematical Methods in Medicine, vol. 2022, 2022C. Milias et al., "Metamaterial-inspired antennas: A review of the state of the art and future design challenges,"IEEE Access, vol. 9, pp. 8946-89865, 2021.