An Efficient and robust ALPR Model using YOLOV8 and LPRNet

1st D.Abdus Subhahan
Computer science Engineering
B V Raju Institute of Technology
Narsapur,Medak,India
subhahan.d@bvrit.ac.in

2nd S. Raga Divya
Computer science Engineering
B V Raju Institute of Technology
Narsapur,Medak,India
21211a05t2@bvrit.ac.in

3rd U. Kavya Sree
Computer science Engineering
B V Raju Institute of Technology
Narsapur,Medak,India
21211a05w3@bvrit.ac.in

4th T. Kiriti
Computer science Engineering
B V Raju Institute of Technology
Narsapur,Medak,India
21211a05v8@bvrit.ac.in

5th Y. Sarthik
Computer science Engineering
B V Raju Institute of Technology
Narsapur,Medak,India
21211a05z1@bvrit.ac.in

***Abstract*—ALPR has many applications in the real world like automatic toll collection, traffic management, and law enforcement. This is generally achieved by two basic steps that are detecting the license plate from the image or video footage and then recognizing the characters on it. In this work, we are going to propose a real-time accurate system using two convolutional neural networks(CNN) that are named YOLOV8 and LPRNet. YOLOV8 is the improved and updated version of YOLO algorithms which has shown improvement in recognition and recall rate and also has better recognition speed compared to other versions with an better accuracy rate.**

***Index Terms*—Deep Learning,License plate detection, license plate recognition, YOLOv8, LPRNet.**

I. Introduction

ALPR has made complex activities like vehicle identification and traffic flow monitoring in the real world simpler using Computer Vision. Since it is having a wide range of applications, researchers have been continuously working on developing a system with better performance and accuracy.

Current traditional ALPR algorithms exhibit well in general scenarios where the performance of the system decreases in difficult environments such as uneven lighting ,oblique views, weather and occulsions. Hence, many models like YOLO, HOG, SSD, R-CNN, and VSNet have been utilising deep learning networks (DNN) due to their potent feature representation. With the increasing population and hectic schedules of people, speed plays a major role. Hence we need a system that can give accurate results in very less time.

Generally, Automatic license plate recognition uses two CNNs for license plate identification and recognition of characters on it, which is a lot of work and time-consuming hence, we use the YOLOV8 algorithm which has great accuracy and robustness. The other versions of YOLO algorithms have some problems like fuzzy license plate character recognition in YOLOV3 and they use anchoring boxes for object identification and it takes lots of time to find the center of the object from the offsets of known anchor box, which means that YOLOV8 algorithm directly predicts the center of the object. The mosaic method means joining the 4 images together and making the model learn objects in new locations ons, used for the training of the model which can be useless to train the model like this during each epoch hence it is advisable to rain the model using mosaic augmentation during the last 10 epoch's.

We are using YOLOV8 as one stage algorithm for license plate detection whose output is then given as input for STN(Spatial Transformers Networks) to correct the detected license plate like slanted, distorted, or any other distortions to give the best quality input to the LPRNet Algorithm for the detection of the characters that are present on the license plate.

II. Literature Survey

The License Plate Recognition (LPR) model based on the improved YOLOv5 algorithm and GRU is a computer vision system that can accurately detect and recognize license plates from images or videos. The system uses a deep learning algorithm called YOLOv5 (You Only Look Once version 5) to perform object detection on the input image or video frames. The YOLOv5 algorithm is an improvement over the previous versions of YOLO and is known for its high accuracy and speed. After detecting the license plate, the LPR system extracts the characters from the plate using a technique called Optical Character Recognition (OCR). A Gated Recurrent Unit (GRU), a type of Recurrent Neural Network (RNN), makes up the OCR module. which is trained on a large dataset of license plate characters. The GRU network takes in the license plate image and produces a sequence of characters that correspond to the license plate number. The LPR system can be used for various applications The system is highly accurate and can recognize license plates even in low-light conditions or when the plate is partially obscured. The improved YOLOv5 and GRU algorithms make the system efficient and robust, allowing it to process large amounts real-time data. The system combines the Improved YOLOv5 and Gated Recurrent Unit (GRU) deep learning models. The Improved YOLOv5 model is a state-of-the-art object identification model that is capable of detecting objects with high accuracy and speed. It is trained on a large dataset of license plate pictures and is used to detect and localize license plates in an input image or video frame. The structure of the YOLOv5 network comprises four components there are input, backbone, neck, and prediction. YOLOv5 has undergone iterations up to version V6.1. In the experiments conducted in this study, we utilized the open-source CCPD dataset, which is a large-scale collection of Chinese urban parking images. Additionally, we included some vehicle data images collected by ourselves, resulting in a total of 12,500 images. All images in the dataset have dimensions of 1160×720 and encompass a wide range of complex environments, such as strong illumination, a foggy sky, dim illumination, smudged license plates, and slanted sceneries. As a result, this collection offers a wide range of samples of license plate scenes. Once the license plate is detected, the extracted license plate image is fed into the GRU model, which is a type of recurrent neural network that is commonly utilised for sequence prediction tasks. The GRU model is developed on a dataset of license plate images and is used to the characters on the licence plate to be able to read and extract. The LPR system based on the Improved YOLOv5 and GRU has several advantages over traditional LPR systems. It is more accurate and faster than traditional systems, and it is capable of recognizing license plates in real time. Additionally, the system can be utilised in a different types of applications, such as traffic monitoring, managing parking systems, and enforcing law. Overall, the LPR system based on GRU and improved YOLO is an effective and efficient solution for license plate recognition, and it has the potential to revolutionize the way license plate information is collected and processed[1].

Analysis of licence plate images using generative adversarial neural networks (GANs) is a research paper that proposes a new approach for license plate recognition using deep learning techniques. The authors of the paper use GANs to generate realistic images of license plates, which are then used to train a recognition system. The approach suggested is assessed on a large dataset of license plate images and achieves state of the art results in terms of speed and accuracy. The paper starts by introducing the problem of license plate recognition, which is a crucial undertaking in many applications such as traffic monitoring, parking management, and law enforcement. The authors then explain the limitations of traditional recognition systems, which rely on handcrafted features and require extensive preprocessing of the images. The proposed approach is based on the utilisation of GANs, which are a type of deep neural network that can generate realistic images. The authors train a GAN to generate license plate images that are similar to real ones and then use these images to train a recognition system using a convolutional neural network(CNN). The CNN is trained on both real and synthetic images, which allows it to learn robust features for recognition. The experiments findings demonstrate that the suggested strategy works. achieves state of the art performance on the several benchmark datasets. In particular, it outperforms previous approaches in terms of accuracy and speed. The authors also conduct a detailed analysis of the results, showing that the proposed approach is robust to various types of noise and distortion. Overall, the paper presents a novel approach to license plate character recognition that is based on the use of GANs. The approach achieves state-of-the-art results and has the potential to improve many applications that rely on license plate recognition[5].

In this research paper, YOLOv5m and LPRNet methods are used for License Plate recognition to overcome the shortcomings faced when methods like YOLOv4-LPRNet, YOLOv5s, YOLOv5- LPRNet are used. In YOLOv5m and LPRNet the accuracy is 99.49 percent and 98.79 percent respectively. YOLOv5m is used to detect license plates from the moving car and LPRNet is used to recognize the characters on the license plate without using character segmentation. The major shortcomings faced In contrast to other approaches, which have limited real-time performance and recognition accuracy, the algorithm utilized in this study is resilient and operates quickly. In YOLOv5m-LPRNet, the car's license plate is utilized as the object, the recognition test is done, and the accuracy of recognition is quantified. The four components of the YOLOv5m network are input, backbone, neck, and prediction. The mosaic method is used in Input to realize data enhancement. Input implements data improvement via the mosaic method. For data training, the mosaic method involves picking arbitrarily numerous images to combine on one image. The image is then flipped and the color gamut is tweaked. Focus, CBS, CSP1-X, and SPP 4 modules are deployed by Backbone to accomplish image feature extraction. Slice the input image is the main concern and which reduces the number of parameters speeds up the convolution process. CBS is made comprised of the activation function, batch normalization layer, and convolution layer. Some of its primary applications on slice images include convolution, normalization, and activation techniques. Two configurations of the Cross Stage Partial (CSP) network exist: CSP1- x and CSP2- x, where x is the quantity of residual components. To easily prevent gradient disappearance, feature extraction employs CSP1-x.MaxPool and CBS make up the Spatial Pyramid Pooling (SPP), expanding a field is its primary purpose. It employs different-sized convolution kernels. The neck section achieves multi-scale feature information fusion. It was made by combining the Feature Pyramid Network (FPN) and the Pyramid Attention Network (PAN). In this instance, the FPN and PAN both perform top-to-bottom upsampling and bottom-to-top downsampling on the feature map, and the output of both networks is fused using the CSP22-2 module. In accordance with the supplied image attributes, prediction determines the boundary box target category and target position with the best confidence score. The three most frequently used bounding box loss functions are location, confidence, and classification loss functions. Without segmentation, LPRNet can complete comprehensive end-to-end license plate recognition by using a lightweight convolutional neural network structure. LPRNet's backbone consists of two dropout layers to avoid overfitting and three layers each of convolution, maximum pooling, and fundamental modules. The 94 x 24 image is the input for this network, while the convolutional layer is the output. And each fundamental module has two input layers and four convolutional layers. Here to avoid errors CTC(Connectionist Temporary Classification) is used[4].

(ALPR) approach is used to get a more accurate output by using processing pf image and recognition of pattern methods, characters on license plate can be extracted and recognized from videos or images. VSNet is a novel ALPR approach which is a cascaded framework that uses two Convolutional Neural Networks that are 1. Vertex Net- To detects the License Plate 2. SCR-Net -To recognize the License plate. Experiments results indicate VSNet achieves better than 99 percent (99) accuracy on the AOLP and CCPD Datasets with 149 FPS inference speed, outperforming state-of-the-art approaches by 50 percent more improvement in error rate. Oblique view, uneven illumination, different weather conditions, and small size can be handled using vertex net designed with small-resolution input but The recognition performance remains unaffected despite the resampling and rectification of license plates (LPs) to higher resolutions based on the predicted vertices by VertexNet. Both VertexNet and SCR-Net are designed with unique architectures to implement vertex-based operations. This system efficiently incorporates vertex estimation and LP rectification using a CNN structure, achieving optimal inference speed and accuracy.Every ALPR System consists of two components license plate detection and recognition. VertexNet: Traditional LP detection method depends on extracting features like texture, colour, etc which could be sensitive to complex backgrounds(objects with the same color and texture).In contrast, CNNs outperform conventional techniques in terms of feature representation and performance. Vertex estimation, small input sizes, and a narrow channel of high-level layers can all be used to fabricate the one-stage detector known as vertex not. It has a detector with a small input size of 256 x 256 compared to YOLO Networks (416 x 416) to detect vehicles. It has 3 modules 1. backbone 2. Fusion and 3.Head Networks. Backbone has 6 stages in each stage the size of the output is half of the input and these outputs are fed to the next sequential stages consisting of Integration Blocks (IBs) which are s building blocks of vertexNet. The fusion network fuses the necessary features of the backbone network. In addition to performing LP classification and box offsets prediction, the head network also includes a vertex-estimation branch. SCRNet: The output obtained by VertexNet is taken as input by SCRNet. It takes it as a classification problem rather than recognition like YOLO-based detection methods and performs forward pass CNN. It first resamples and rectifies the output from VertexNet based on VertexNet's anticipated vertices. Then the weight-sharing classifier addresses the issue of smallscale datasets as it spots all the instances of the characters in datasets and classifies based on the probability of each character. VertexNet improves precision by 1.4 percent compared to state-of-the-art MTLPR obtaining the highest accuracy 99.1 percent and runs at a speed of 5.7ms per image with an accuracy of 98.8 percent it is also more effective than cutting-edge techniques. SCR-Net has the highest accuracy (99.5%). on overall testing sets with a speed of 6.7ms per image which is a 43 percent improvement compared to RPNet of 11.7ms per image. There is only a 0.5% error rate[3].

III. Problem Statement

- The previous state-of-art models take large input sizes and process them using technologies like CRNet that use the segmentation method for character recognition which is a laborious task that results in low inference speed and does not perform well with small-scale data sets.
- • Their accuracy decreases in unconstrained environments such as occlusions, bad weather, and oblique view as the image or video quality obtained will be of low resolution.

IV. Objectives

1) Automatic license plate recognition using YOLOV8 and LPRNet.
2) Robust lightweight LPRNet framework to produce accurate results in less time.
3) Detecting fake license plate

V. Existing Work

Machine learning and Deep learning play a major role in the detection of the license for the development of intelligent transport systems. The vehicle number plate recognition process is divided into two stages the identification of the vehicle number plate from the image or real-time video footage and the identification of the characters present on it. The detection of the license plate location from the vehicle is carried out by using machine learning models like Haar-cascade and deep learning architectures like Single Shot MultiBox Detector(SSD), Histogram of oriented gradients(HOG), R-CNN(Region-Based Convolutional Neural Networks), YOLO algorithms and text extraction or character recognition from the detected number plate is carried out using algorithms like LPRNet, Connectionist temporal classification (CTC), Gated Recurrent Units(GRU)[1], Generative Adversarial Neural Networks (GANs).SSD Algorithm uses a single deep neural network to locate license plates in the images or footage. To extract the features, HOG creates histograms based on the gradient's amplitude and directions. For better object detection, R-CNN is used to localize the items in an image based on their pixel intensities. To identify bounding boxes and forecast object class probabilities, YOLO suggests using end-to-end neural networks. LPRNet[4] is used to recognize the characters on the license plate without segmentation. GRU+CTC[5] are used to shorten the training time and increase the convergence speed. GANs[5] are used because they are capable of generating realistic super-resolution images.

VI. Proposed Work

YOLOv8 (you only look once) algorithm is the newest version of the YOLO real-time object detection model and is developed by ULTRALYTICS. The features that make YOLOv8 stand out from the other versions of YOLO are its great recognition speed and accuracy which enhances its performance and flexibility. The network structure of YOLOv8 is a bit complex compared to the previous models but the model is quite efficient. The architecture consists of Backbone, Neck, and Head.

Backbone is a series of Convolutional layers performing or implementing image feature extraction, which includes CSP1_x and Focus, four modules of Spp, and CBS. Fragmenting the source image, which results in a smaller feature image, fewer layers and parameters, and faster convolution processing, is the primary function of Focus. The convolution layer (Convolution), activation function (SiLU), and batch normalization layer (BN) are the three components that encapsulate CBS. Convolution, normalization, and activation processes are what it does to slice pictures. The Cross Stage Partial (CSP) network is made up of two structures, CSP1_x and CSP2_x, where x is the quantity of unaccounted-for components. When features are extracted from the network, CSP1_x with residual structure is used to increase the gradient. As a result, gradient disappearance caused by network deepening is prevented during layer-by-layer back-propagation. The composition of CSP is shown in the following diagram. The SPP module's ability to increase the receptive field is its most significant function. The SPP does not employ a pool operation; rather, it combines various results with the data in order to maintain the same dimension for the output characteristic vector. It does this by using different-sized convolution kernels to input different-sized characteristic maps in order to achieve maximum pooling.

The inclusion of the neck section enables the integration of data pertaining to multiple scale attributes. The structure employed for this purpose is the Feature Pyramid Network (FPN) + Pyramid Attention Network (PAN). which consists of both top-down and bottom-up routes. FPN extracts feature from the feature map by sampling them from top to bottom, combining those features with those from the backbone network; PAN extracts the features from the feature map by sampling them from bottom to top and then combines them with those from the FPN layer. The FPN+PAN network structure is used to combine the features generated from the backbone and detection network, which also enhances the network's capacity for feature fusion. Through the CSP2_2 module, the extracted features are combined. A channel splicing module called Concat is used to realize an assembly of image features. The head portion performs the detection based on the loss metrics using the output of the neck section as input. There are three losses in the conventional YOLO algorithm: class loss, box loss, and object loss.

An anchor-free model is YOLOv8. In other words, rather than predicting an object's offset from a known anchor box, it predicts the object's center directly. Since they may represent the distribution of the boxes from the target benchmark but not the distribution of the custom dataset, anchor boxes were a notoriously difficult component of older YOLO models. To speed up Non-Maximum Suppression (NMS), a challenging post-processing phase that sorts through candidate detections following inference, anchor-free detection decreases the number of box predictions.

Character segmentation and character non-segmentation are the two categories into which license plate recognition techniques can be separated. In the approach based on character segmentation, a license plate image is often segmented using edge detection, template matching, and other techniques, and the characters are then classified using CNNs (Convolutional Neural Networks). However, this approach is highly susceptible to errors and is significantly influenced by outside factors like lighting. This network chooses a lightweight CNN structure that can achieve complete license plate character identification with character non-segmentation. The algorithm can be used to identify license plate photographs whether they are blurry, distorted, bad weather-related, or under other circumstances. This section used this to develop the algorithm for recognizing license plates. The license plate scenario required a lightweight character recognition network called LPRNet with a minimal character tag set and short string length. At the input end, LPRNet processes input impressions using STN. Although STN processing is an optional step, its existence gives us the ability to finish the image's transformation and produce an ideal input image for license plate recognition. Following input and 3x3 average pooling, the picture is transferred to two convolution layers for convolution. The two convolution feature maps are then joined through a concatenation operation to produce features of various sizes. By avoiding over-fitting of the convolution layers with 50 percent dropout, the effect is enhanced. It is then output to the following module via two complete connection layers. LPRNet's core utilizes CNN to extract picture features. The raw RGB image is sent into the backbone, which also has 3 convolution layers,pooling layers, and tiny basic block modules. There are a total of 15 convolution layers and 3 pooling layers, with 4 convolution layers per small basic block module. Regularization is done using the Dropout optimization technique to avoid network over-fitting. The backbone's output can be envisioned as a sequence reflecting the likelihood of the matching character, whose length is correlated with the input image's width. The output result is expressed as a sequence of character probabilities, the last layer of the backbone uses a 1x13 convolution kernel, whose length correlates with the width of the input Feature Map. The small basic block module uses a 1x1 convolution to change the dimension so that the depth of the feature map matches the number of character classes. Since character recognition is implemented by LPRNet without image segmentation, training is done using the CTC (Connectionist Temporal Classification) loss function alone. LPRNet uses two techniques for the inference stage decoding: beam search and greedy search. By using a locally optimal approach and the maximum class probability at each place, greedy search anticipates producing a global optimal result. Greedy Search's probability evaluation is improved by Beam Search. Each search retains the top n maximum probability combinations, from which the final output sequence is derived by computing the global maximum probability. The first successful matching template set based on the motor vehicle license plate standard is returned by LPRNet after the post-filtering stage, where it employs Beam Search to find the top n sequences with the highest probability.

VII. Methods

A. Dataset Description

We have collected a data set from Roboflow named a license plate for training the model for license plate detection. It consists of a total of 1680 images divided into two sets called to train and valid each consisting of 1530 and 150 images. The pictures of the license plate in this data set are from different angles and of different sizes which helps in the efficient training of the model. For LPR we need

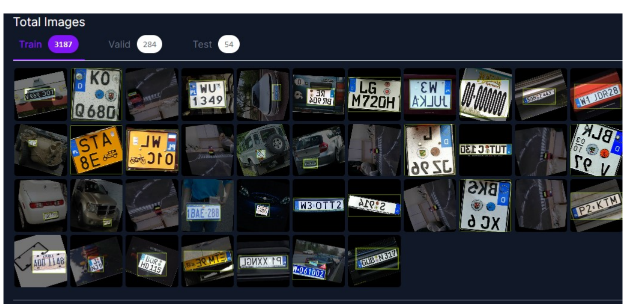

Fig. 1. CCPD Dataset Distribution

to train the model using the dataset consisting of license plates of the same country or region to avoid inaccuracy as different countries have different license plate formats. Hence we are choosing CCPD(Chinese city parking dataset) which also consists of different types like blurr, illuminated, tilted, rotated, and challenging license plate images for the efficient training of the model to get more accurate results.

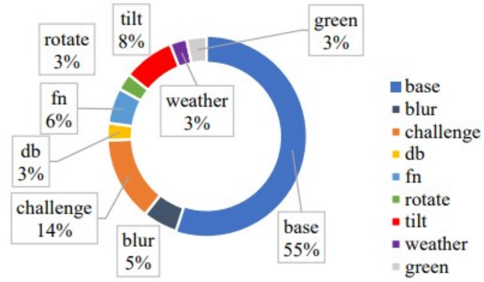

Fig. 2. License plate Dataset from roboflow.

B. License Plate Detection

We train the model using the collected dataset by importing modules called paralytics and basic code from Git Hub, after the training the model stores the results and will be ready to detect the license plate from any input image or video. Firstly, the input image of different sizes and pixels is given to the backbone of YOLOv8, feature extraction is done there inside the convolution layers. The output of this backbone is sent to the neck where the concatenation of all the features extracted undergoes followed by the head where the detection of the license plate is done based on loss metrics.

Fig. 3. Detected License plate with bounding Box

C. Data Preprocessing

The output obtained from the YOLOV8 algorithm is given as an input for the LPRNet algorithm. Before that the output obtained from YOLOV8 must be pre-processed to give the best quality input image for the license plate recognition. To achieve this LPRNet uses STN(Spatial Transformer Network) and reduces the deformations of input images.

D. License Plate Recognition

Once the License plate is detected and processed using STN for deformations, the characters on the license plate must be recognized. To achieve this we use the LPRNet algorithm. The output obtained from STN is given as input for the backbone network of LPRNet which uses CNN to extract the features of the image and applies a kernel that gives the output result of a sequence showing the probability of the corresponding character. In the nest stage, it uses Beam Search to decode the sequence and finds the sequences with the highest probability based on which the characters of the license plate are recognized.

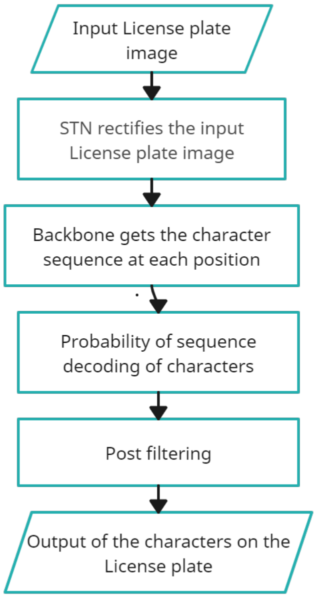

Fig. 4. steps of the LPRNet algorithm

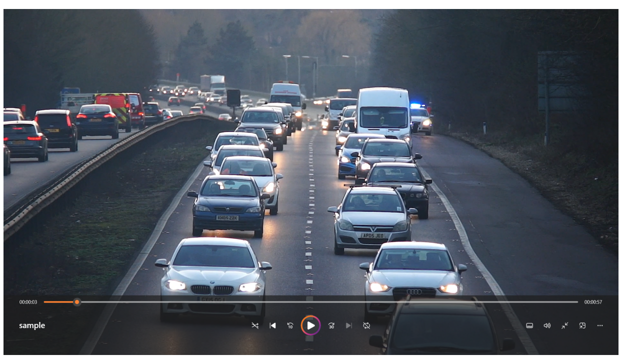

Fig. 5. sample input

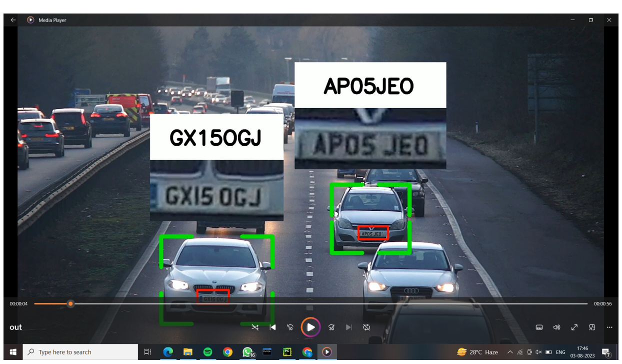

Fig. 6. output obtained

VIII. Result Analysis

The performance and efficiency of this model is measured using the factors Recal, precision, Fscore, Map(mean of average precision), and FPS(frame per seconds).and the higher quantity, faster the activity. The nearer Map value is one of the finer general effectiveness of the model. The evaluation index can be calculated using the equation which uses the confusion matrix is

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\%$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\%$$

$$\text{Fscore} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^{N} \int_{R_1}^{0} P_i(R)\, dR$$

$$\text{FPS} = \frac{\text{Figure Number}}{\text{Total Time}}$$

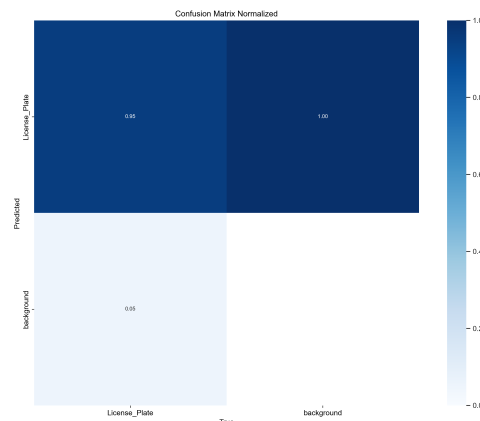

Fig. 7. Confusion matrix

Model	Precision (%)	Recall (%)	F Score	mAP (%)	FPS (figure/s)
YOLOv3-LPRNet	87.46	82.56	0.8494	88.31	30
YOLOv4-LPRNet	89.23	84.31	0.8670	89.21	34
YOLOv5s-LPRNet	93.36	94.86	0.9410	93.51	40
YOLOv5m-LPRNet	94.49	95.37	0.9493	94.59	37
YOLOv8-LPRNet expected results	98	97	0.9714	97	42

TABLE I
Performance comparison of different license plate recognition models

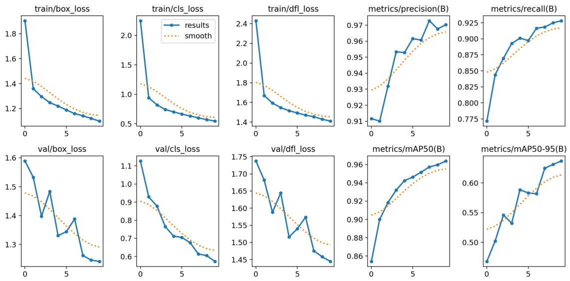

Fig. 8. loss metrics

IX. Conclusion

The YOLOv8 algorithm is the improved version of all the previous YOLO models and is free of anchor boxes and has faster recognition speed and great performance compared to all other models. The LPRNet is used for the recognition of the characters on the detected license plate. The car license plate identification model experiment showcases impressive results in terms of recognition accuracy,recall,and Map(mean average precision).have all improved, reaching 98.49%, 97.9%, and 98.56%, respectively. This method has high accuracy and it can essentially fulfill the requirements for car license plate identification in complicated settings.

References

[1] Henglang, Shi,Dongnan Zhao “License plate Recognition System Based on Improved YOLOv5 and GRU” , IEEE Access vol-11., 30 January 2023.

[2] Qi Wang, Xiaocheng Lu, Cong Zhang “Lsv-lp: large-scale video-based license plate detection and recognition” , IEEE Transactions On Pattern Analysis And Machine Intelligence, vol 45,no 1 , 1 January ,2023.

[3] Yi wang, zhen peng bian, yunhao zhou, lap pui chau “Rethinking and designing a high performing automatic license plate recognition approach” , IEEE Transactions on Intelligent Transportation Systems, vol 23, July 2022.

[4] Shan Luo, Jihong Liu “Research on Car License Plate Recognition Based on Improved YOLOv5m and LPRNet” , IEEE Access vol-10 , 1 September 2022.

[5] Mustafa A. Elattar, Ibrahim H.EL-SHAL, Omar M. Fahmy, “License Plate Image Analysis Empowered by Generative Adversarial Neural Networks(GANs)” , IEEE Access vol-10 , 8 March 2022.

[6] Tabarbour, Amman, Jordan “Real Time Jordanian License Plate Recognition Using Deep Learning “ , Journal of King Saudi University-Computer and Information Sciences 34 , 2022.

[7] Shenghu Pan, Jain Liu, Dekun Chen “Research and License Plate Detection and Recognition System Based on YOLOv7 and LPRNET” , Academic Journal of Science and Technology , vol.4 , No .2,2022.

[8] Thavavel Vaiyapuri, Scachi Nandan Monhanty,Irina, M.Sivaram “Automatic vehicle license plate recognition using optimal deep learning model” , Computers,Materials and Continua, vol 67 ,2021

[9] Imran Shafi, Imtiaz Hussain, Jamil Ahmad “License Plate Identification And Recognition in aNon Standard Environment Using Neural Pattern Matching” , Springer , 2021.

[10] Sergio M. Silva, Claudio Rosito Jung “A Flexible Approach For Automatic License Plate Recognition In Unconstrained Scenarios” , IEEE Transactions on Intelligent Transportation System , 19 June 2021.