



Enhancing ALPR: a two stage YOLO model with data augmentation for improved accuracy and robustness

Swati Bansal¹ · Abhilasha Jain¹ · Manoj Sharma² · Gautam Kumar¹ · Shivam Ojha¹ · Hemant Walia¹

Received: 26 July 2023 / Revised: 14 March 2024 / Accepted: 15 May 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Automatic licence plate detection and recognition (ALPR) systems becomes an important task for a variety of applications such as traffic control, toll collection, law enforcement, and parking management. A growing number of innovative licence plate identification systems are now capturing the interest of scientists worldwide. Therefore, ALPR has become a popular study subject for researchers. However, a large portion of the continuing solutions, which often rely on many imperatives, are still not resilient in the current environment. In this paper, an ALPR system based on the cutting-edge YOLO object detector has been presented. YOLO model is trained and fine-tuned for each ALPR stage to ensure robustness under different conditions such as variations in camera, lighting, weather conditions and background. Especially for character segmentation and recognition, we design a two-stage approach employing double data augmentation tricks such as inverted License Plates (LPs) and flipped characters. Noise is also added to the images to improve the confidence level. The most popular datasets and benchmarks are used to train and evaluate the proposed ALPR system. The recall and accuracy attained for the proposed system is 80% and 94% respectively. our proposed system is well suited to improving safety and efficiency in transportation and law enforcement.

Keywords Object detection · YOLO architecture · Licence plate recognition · Machine learning · Accuracy · Deep learning

✉ Abhilasha Jain
abhilashagzscet@gmail.com

✉ Manoj Sharma
neelmanoj@gmail.com

¹ Department of C.S.E, Giani Zail Singh Campus College of Engineering & Technology, MRSPTU, Bathinda, Punjab, India

² Department of ECE, Giani Zail Singh Campus College of Engineering & Technology, MRSPTU, Bathinda, Punjab, India

1 Introduction

Some of the biggest issues in future smart cities are traffic safety and congestion. Any assistance in running these cities would be helpful to everyone. More than ever, effective automatic licence plate recognition systems are needed to support the growth of smart cities and, by extension, intelligent transportation systems. Due to their use in intelligent surveillance systems with numerous use cases, including automated parking lot management, traffic surveillance, vehicular access control, etc., which represent developing research fields within the context of urban mobility, these systems have attracted a lot of attention. Devices connected to the Internet of Things (IoT) are the foundation of smart cities. These are tiny digital gadgets that are widely employed to carry out data processing activities. Despite being more powerful than regular IoT devices, IoT processing devices like Raspberry Pi, Google Coral, NVIDIA Jetson Nano, etc., lack constant processing and cooling abilities, in contrast to conventional machine learning workstations. In order to help the evolution of smart cities, this study seeks to show a system that has been built that is both time and size efficient and can be installed on tiny IoT processing devices. ALPR systems are useful in these types of situations. ALPR systems typically have three stages: License Plate (LP) detection, character segmentation and character recognition.

Even though ALPR has been discussed in the literature extensively, many research and approaches are still insufficiently reliable in real-world situations. These methods frequently rely on certain restrictions, such as certain cameras or viewing positions, plain backdrops, ideal illumination, search in a specified area, and specific kinds of automobiles (they wouldn't be capable of recognizing LPs emerging from motorcycles, lorries, or buses in moving condition).

Deep Learning (DL) techniques are used in this situation. However, there is still a high interest for ALPR datasets with cars and LPs descriptions despite the impressive advancement of DL techniques in ALPR. The effectiveness of DL approaches depends on the quantity of training data [1]. In the present manuscript an ALPR model is proposed and main highlights of the manuscript are:

- 1 The proposed ALPR system detects the LP precisely in real time using images and streams of videos.
- 2 The proposed model detects LPs of different vehicles like cars, motors cycles with different aspect ratio and positions. Moreover, the system works to detect LP of vehicles of different countries like India, China, UK, America and so on.
- 3 To train proposed model, different types of vehicles like cars, motorcycles, buses and so on, camera positions, camera elevations and different climate conditions incorporated in training data set to meet the real time scenarios. Therefore, our proposed model deals with two challenges. Firstly, partial visible images due to different weather conditions. Secondly, car and motorcycle LPs have different aspect ratio, layout, and positions.
- 4 Despite the number of ALPR systems under progress, the majority of these significantly rely on powerful computing resources. In this research, we offer a system that is considerably more compact and suitable for low-powered computer devices.
- 5 There are two subsystems in the proposed strategy. The first component of the suggested system, used to identify licence plates, is based on YOLOv5. Secondly, we utilised the open-source OCR engine commonly name as EasyOCR for character identification.

The structure of the manuscript is as follows: Section 2 presents the literature survey in the field of ALPR and Section 3 presents the proposed methodology and Section 4 comprises experimental analysis. Finally, Section 5 concludes the paper.

2 Literature survey

Every country now has a serious issue with traffic regulation and vehicle ownership identification. It might be challenging to recognise vehicle owners who drive excessively fast and against the regulations of the road. As a result, it is impossible to apprehend and penalise those individuals since traffic officials may not be able to obtain the car's licence plate from a moving vehicle due to its speed. As one of the answers to this problem, it is necessary to design an Automatic Number Plate Recognition (ANPR) system. Several ANPR systems are proposed by researchers but still it is a challenging task to recognise the number plate as several elements, such as a vehicle's rapid speed, non-uniform number plate, language of the number, and changing lighting circumstances, can have a significant impact on the overall identification rate. Majority of the systems function with these restrictions. Using picture size, success rate, and processing time as factors, ANPR techniques are presented [2]. Algorithms for licence plates recognising (LPR) in pictures or videos typically involve the following three phases in the processing chain: First, a licence plate region is extracted, next the characters on the plate are segmented, and finally, each character is recognised. Due to the variety of plate designs and the uneven outside lighting conditions during image collecting, this work is rather difficult. Therefore, in this segment, a brief overview of several recent works that apply DL methodologies to ALPR is presented.

Anagnostopoulos et al. [3] reviewed, categorised and evaluated LPR approaches that have been developed for still photos and video sequences, when accessible, issues like processing speed, computing capacity, and recognition rate.

An effective approach for vehicle license plate detection and recognition, based on character-specific regions has been presented on static images in [4]. Initially a sequence of morphological operations is applied to find plate candidates with dense vertical edges. Then the character-specific extremal (ERs) are extracted and selected as character regions in color space. The recognition step is achieved by an effective hybrid discriminative restricted Boltzmann machine (HDRBM) classifier. The proposed method achieved 95.9%, 98.2%, 91.9% and 94.1% better average performance with LDR, CRR, OVR1 and OVR2 respectively.

In another study proposed by Hsu et al. 2013 [5] suggested three main categories of vehicle LPR applications and suggest a solution with parameter settings that may be altered for various applications. Three selected categories were Road patrol (RP), access control (AC) and law enforcement. In this paper, plate detection is solved using edge clustering that performed better than many earlier solutions.

Ashtari et al. [6] proposed a hybrid classifier (SVM and a simple decision tree) that can recognise licence plate characters and an Iranian car licence plate recognition system based on a novel localization approach is presented. The proposed approach employed an altered template-matching algorithm to analyse target colour pixels in order to locate a licence plate. The typical color-geometric template used in Iran and numerous European nations can be localised using a modified strip search. The system's overall performance attained is 94%, with a 96% performance detection rate.

Sarfraz, M.S., et al. [7] offered a powerful real-time ALPR framework that is especially made to operate on CCTV video footage received from cameras that have not been specifically set up for ALPR. The suggested method has the unique ability to automatically adjust for different camera distances and lighting conditions, which is necessary for a video forensic tool that can work with movies taken by a variety of unnamed, scattered CCTV cameras.

Panahi et al. [8], proposed a highly accurate online ANPR system that can serve as the foundation for a variety of applications. Their proposed model is able to handle illegible licence plates, changing weather and lighting conditions, various traffic scenarios, and fast-moving cars. The authors offer appropriate hardware platforms, real-time, reliable, and creative algorithms, and tackles many challenges. Additionally, this system was tested on three additional Iranian data sets, and it was 100% accurate in both the recognition and detection portions.

In order to precisely localise car licence plates from complicated situations in real time, Yuan et al. [9], provided a reliable and effective method for licence plate detection. In order to distinguish the real licence plate from the candidate regions, a cascaded licence plate classifier utilizing linear support vector machines and leveraging colour saliency features is proposed. Montazzoli, S. and C. Jung [10] suggested an acceptable Convolutional Neural Network (CNN) architecture-based end-to-end DL-ALPR system for Brazilian licence plates. Azam and Islam [11] introduced a brand-new ALPD technique that can efficiently identify LP areas from images under risky circumstances. A unique technique was used to remove streaks of rain from images by using a frequency domain mask. The suggested rain removal method outperformed the single-image based rain removal method currently in use. For managing low intensity indoor, night, blurry, and foggy pictures, a unique contrast enhancement methodology using a statistical binarization approach is included in the suggested ALPD approach [11]. Rafique, M.A et al. [12] proposed a methodology using cutting-edge object detection approaches, such as convolutional neural networks with region proposal (RCNN), its latest mode (Fast-RCNN and Faster-RCNN), and the exemplar-SVM. The proposed study showed superior outcomes over other traditional methods in thorough tests and comparisons. Svoboda, P., et al. [13] investigated the previously described method of direct blind deconvolution and noise removal with convolutional neural networks (CNN) in a case where the blur kernels are somewhat limited.

Based on the cutting-edge YOLO object detector, the ALPR system presented is reliable and effective by Laroca, R., et al. [14]. For each ALPR step, the Convolutional Neural Networks (CNNs) are developed and adjusted such that they are resilient under various circumstances (e.g., difference in camera, variable lighting conditions, and background). A two-stage method specifically for text segmentation and recognition using low-tech data augmentation techniques like reversed licence plates (LPs) and inverted characters have been created. After the YOLO series failed to gain traction, G. Yang et al. [15] released a paper, continued to improve it, and eventually won YOLO's official endorsement in 2020. YOLOV5 can reach 140 FPS on the Tesla P100 quick detection, but YOLOV4 can only manage 50 FPS. YOLOV5 is merely 27 MB in size, however YOLOV4 employing Darknet architecture is 244 MB large. The accuracy of YOLOV5 is identical to that of YOLOV4. YOLOV5 has carried over the benefits of YOLOV4, including the addition of SPP-NET [15].

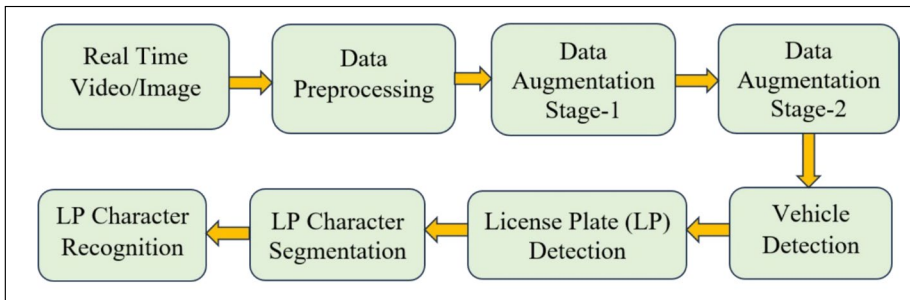
3 Proposed methodology

Herein, a licence plate detection methodology using YOLO5 is proposed. The proposed methodology is able to detect the licence plate of a vehicle from static images (jpeg, png, webp etc.) as well as from real time videos. The proposed model is able to detect the licence plates of many countries like India, UK, US and EU. Different lighting conditions and weathers have been considered to achieve the objectives of the paper. Figure 1a shows the proposed methodology and the proposed approach runs with various phases as: Data pre-processing, two stage data augmentation, Vehicle detection, Licence Plate detection and plate number recognition. The following subsections presents the dataset description and the implementation of various phases.

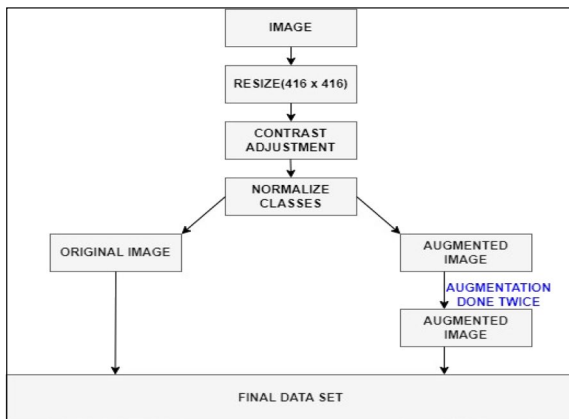
3.1 Dataset description and pre-processing

The original dataset of licence plate is collected from following open-source platforms and are available publicly:

- (a) <https://www.kaggle.com/datasets/andrewmvd/car-plate-detection> [19];
- (b) <https://www.kaggle.com/datasets/thamizhsterio/indian-license-plates> [20] and;



(a)



(b)

Fig. 1 a Proposed Methodology for ALPR Model. b Data pre-processing and image augmentation



Fig. 2 Augmentation Steps applied on original image

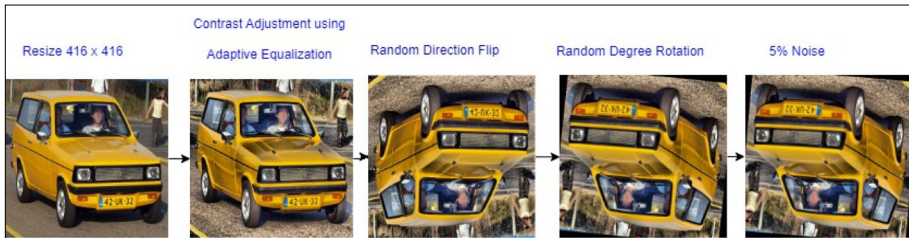


Fig. 3 Image after augmentation

(c) <https://universe.roboflow.com/samrat-sahoo/license-plates-f8vsn> [21].

Pre-processing steps are applied to the collected raw data set shown in Fig. 1b.

The collected image data set is labelled where bounding box coordinates of objects in the image are also given. The collected images are in different formats like PNG, JPEG, WEBP etc. which then converted to PNG formats of size 416×416 using Roboflow software. The reason to take size of image 416×416 is that the real time videos are best suited with this image size and format. Moreover, training time for this format and size is optimal and training of the model use less resources.

Now the curated data set contained 804 images of size 416×416 in PNG format. Contrast adjustment of the images has been done to sharpen the image colours and the normalization the image classes (LP, LICENCE, LICENCE PLATE, PLATE) into Licence Plate as data is collected from different open sources. Two step augmentation as shown in Fig. 1 is performed on original data set to have greater spectrum of semantic variations in the training data set and to achieve generality of model. This increase in training samples also evades overfitting of model. The series of steps performed for augmentation are presented in Fig. 2. Random flip is added in augmentation to consider all the scenarios of camera angle in natural scenes. 5% noise is added through bounding pixelation to simulate distortions in images due to climate change. After repeating two step augmentation, new data set has 2412 images. In Fig. 3, some samples of augmented data set are shown.

3.2 Vehicle detection

The first phase detects the coordinates of bounding boxes to find the location of vehicle in the scene. To choose an object detection model, factors like high accuracy rate, high calculation speed and calculation cost of model are taken in consideration. The model needs to have high accuracy rate as any missed vehicle will result in an overall failure to detection of licence plate system. The selected model should perform calculations efficiently so that the proposed system can deploy in real time scenarios. After careful consideration, we have decided to use YOLOv5 transfer learning model for object detection. The architecture of YOLO5 is presented in Fig. 4. It comprises of three parts: backbone, neck and head. Backbone comprises a hybrid CNN known as CSPDarknet

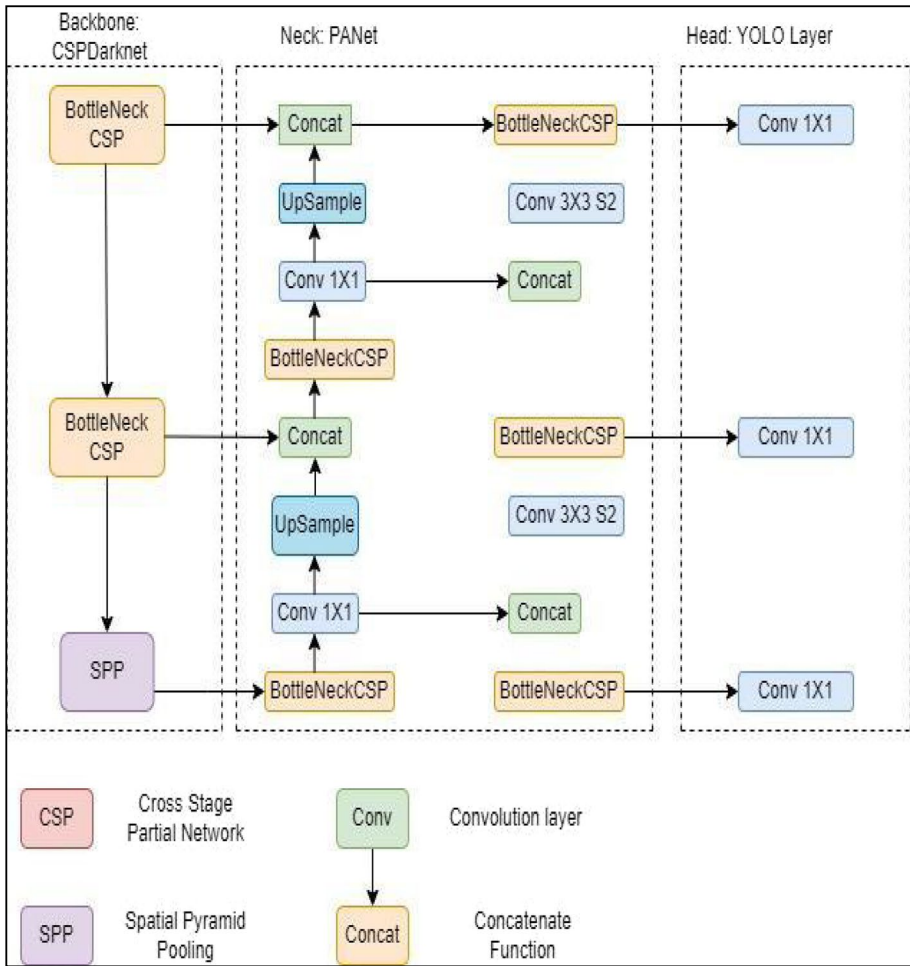


Fig. 4 Architecture of YOLO5

which integrates the CSPNet with Darknet. CSPNet resolves the glitches of repeated gradient in YOLO4 by saving gradient changes in the feature map and reduces the size of model. It also improves the speed and accuracy of model which makes model suitable for LP detection in real time scenarios. Neck of YOLOv5 is path aggregation network (PANet) which is a sequence of layers that augment the location accuracy of the object in image features and generate feature pyramids of 3 different sizes (18×18 , 36×36 , 72×72). These pyramids enable the model to scale objects in general and identify same object with different sizes and scale. The final detecting stage is primarily carried out using the model Head. This part applies anchor boxes on features. A final output vector with class probabilities, object scores, and bounding boxes has been produced as shown in Fig. 5. As shown in Fig. 4, in YOLO's CSP backbone, "bottleneck" refers to a design aimed at reducing computational complexity and preserve performance. CSPNet, a common variant of the CSP backbone used in YOLOv5, uses a bottleneck layer to reduce feature map dimensionality before processing.



Fig. 5 Final output vector with class probabilities, object scores, and bounding boxes

3.3 Licence plate detection

The second phase detects the licence plate of detected vehicle. As it has been found that the format and colour of licence plate varies from country to country for example yellow colour licence plates are assigned to vehicles in Brazil whereas blue and yellow colour plates are assigned light and heavy vehicles in China. In countries like India, US, EU and UK white colour licence plates are assigned with standard black characters. So, the proposed phase offers a framework to recognize LPs of different nations with variety of colours for characters and backgrounds.

To detect LP in the detected vehicle, a custom YOLOv5 model runs over frames of detected vehicle in the video. The algorithm divides the frame into N number of grids with equal size. Each of these N grids is responsible for the detection and localization of LP in the frame of detected object. Correspondingly, these grids predict B bounding box coordinates relative to their cell coordinates, along with the object label and probability

of the object being present in the cell. This process greatly lowers the computation as both detection and recognition. It brings forth a lot of duplicate predictions due to multiple cells predicting the same object with different bounding box predictions. YOLO makes use of Non-Maximal Suppression to deal with this issue. In Non-Maximal Suppression, algorithm suppresses all bounding boxes that have lower probability scores.

YOLO achieves this by first looking at the probability scores associated with each decision and taking the largest one. Following this, it suppresses the bounding boxes having the largest intersection over Union with the current high probability bounding box. This step is repeated till the final bounding boxes of license plate are obtained.

3.4 Character segmentation

After running the Custom YOLOv5 model over an image, the coordinates of detected LP are obtained. Firstly, the image of LP with the detected coordinates is cropped from the frame and colour scheme of cropped image is changed from RGB to Grayscale. Because each country has its own color plate so to standardised into one color, it should be converted into Grayscale. Secondly, an artificial image (mask) of #000000 colour (total black) of same size as the cropped bounding box image has been created and superimposed on the cropped bounding box image. This step keeps only the black characters of license plate and removes everything else from the image. Thirdly, Canny Edge Detection Algorithm has been applied to separate each character images from the cropped image. The edge detection algorithm finds the edges (contours) of the characters. Then, the contours are sorted by area size. Iterate over each contour and subtract #000000 (total black) image from the contour. This step makes sure we are getting the character part only and removing the non-character part if by any chance it appears in contour. Character will turn white after this step. This method is used to remove white noise. Finally, white plate characters with black background will be obtained. Table 1 presents the algorithm for number plate character segmentation and Fig. 6 shows the visual representation of each step of proposed algorithm presented in Table 1.

Table 1 Algorithm for number plate character segmentation

1. *Crop Detected License Plate.*
2. *Convert RGB to Grayscale.*
3. *Remove noise by using noise erosion and image dilation.*
4. *Create black image mask.*
5. *Superimpose mask on processed image.*
6. *Separate each character using the below algorithm:*
 - (a) *Find all contours using Canny edge detection algorithm.*
 - (b) *Sort the contours by area size.*
 - (c) *Iterate over all contours and subtract black from contour.*
 - (d) *Remove white noise from character contour.*
 - (e) *Save character as separate image.*
7. *White Plate Characters with Black BG will be obtained.*

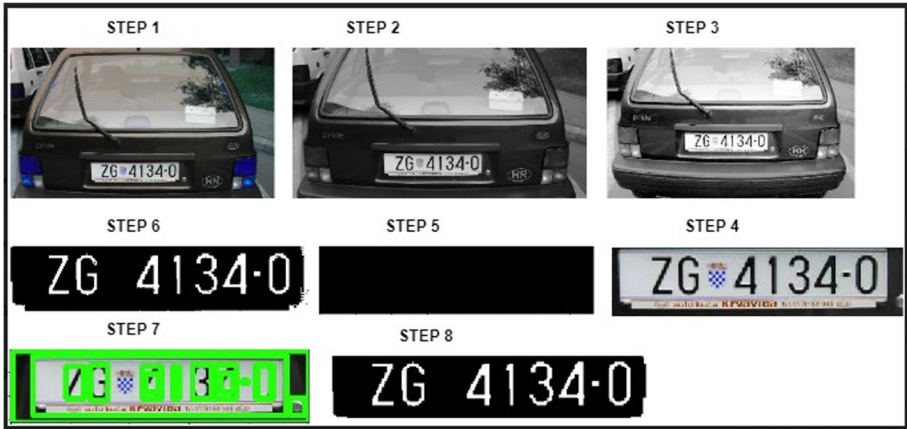


Fig. 6 Visual representation of Algorithm presented in Table 1

3.5 Character recognition

EasyOCR has been used for character recognition in the proposed work. EasyOCR is a python package that can read more than 42 languages all over the world and performs well for natural scene images. This package is used to process digital images and outputs a normal editable text and searchable data. The framework of Easy OCR is shown in Fig. 7. EasyOCR utilizes the **CRAFT Model** for scene text detection and afterward gives the detected text mAPs to the custom **CRNN model** for text recognition. It consists of **3 principal components** namely - *feature extraction* (using Resnet and VGG), *sequence labeling* (using LSTM) and *decoding* (using CTC). Output after OCR have shown in Fig. 8 and the proposed scheme offers following advantages:

1. Although the object (Car) is at quite a far-away distance from the camera, our model was able to accurately locate the target (LP), and perfectly process the OCR.

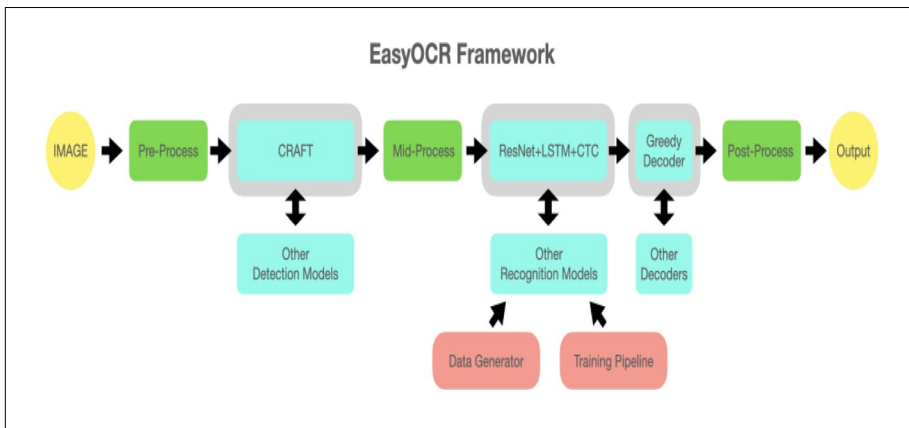


Fig. 7 Proposed framework for character recognition using EasyOCR



Fig. 8 Detected number plate outputs at different conditions using EassyOCR

2. Though the object (Car) is present in skewed angle, our model again accurately detected the target (LP) and perfectly processed the OCR.
3. Even though, in this image the conditions to read the plate text are not suitable even for some human eyes. Our model was able to do so accurately.

4 Experimental analyses

The training and testing of number plate detection and identification were simulated in Google Collaboratory. For all the experiments, the data source is arbitrarily split into training and test datasets. The training data source (80%) is employed for building and training of the model and the test dataset (20%) is engaged for testing the model.

4.1 Principle of the detection algorithm

Network Module for YOLOv5 Glenn Jocher in 2020 [16] introduced the one-stage target recognition method known as YOLOv5. YOLOv5 may be separated into four hierarchical network variants: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, based on variations in network depth and breadth. The YOLOv5s network among them has the highest computation speed but the lowest average precision, while the YOLOv5x network exhibits the opposite traits. The YOLOv5 network's model size is roughly one-tenth those of YOLOv4 network. Its accuracy is on par with YOLOv4 and boasts quicker identification and placement rates. The Backbone, Neck, and Head make up the three primary parts of the YOLOv5 network. Backbone gathers and creates picture characteristics based on various image granularities when the image is entered. Then, Head forecasts the image characteristics to provide bounding boxes and anticipated categories after Neck stitches the image information and sends it to the prediction layer [17].

A state-of-the-art detector dubbed TPH-YOLOv5 is created by combining several cutting-edge methods, such as transformer encoder block, CBAM, and some expert tricks with YOLOv5, and is particularly effective at detecting objects in drone-captured situations.

Author demonstrated that TPH-YOLOv5 obtained cutting-edge efficiency in VisDrone2021 dataset using record of VisDrone2021 dataset. They experimented with a lot of characteristics and utilized some of them to increase the object detector's accuracy [18].

4.2 Results

Results of proposed model ALPR with single augmentation and after two step augmentation on dataset have been shown in Figs. 9, 10 and 11. From Fig. 9a-b, it has been observed that with double augmentation, the proposed model can detect multiple licence plates easily and accurately as compared to single augmentation. The confidence level achieved with double augmentation model is more for front LP of vehicle as compared to single augmentation model as shown in Fig. 10a-b. From Fig. 11a-b, it is evident that even in the ideal scenario (back facing car) double augmentation model is more confident about the detection of LP. The reason, we use augmentation in pre-processing to introduce diversity in the data (mutate the samples). Now what double augmentation does is it takes the already diverse dataset and mutates it even more at such higher level that when we train the model over it, we get all kinds of situations in the data, thus causing good results.

The proposed model has implemented in Python. The performance of the proposed model is evaluated in terms of Precision, Recall, and mean average precision and are represented in Figs. 12, 13, 14, 15 and 16. Table 2 presents the results in terms of precision, recall and mAP with IoU threshold of 0.5. The IoU is a measure of overlap between the predicted bounding box and the ground truth bounding box. An IoU of 0.5 means that the

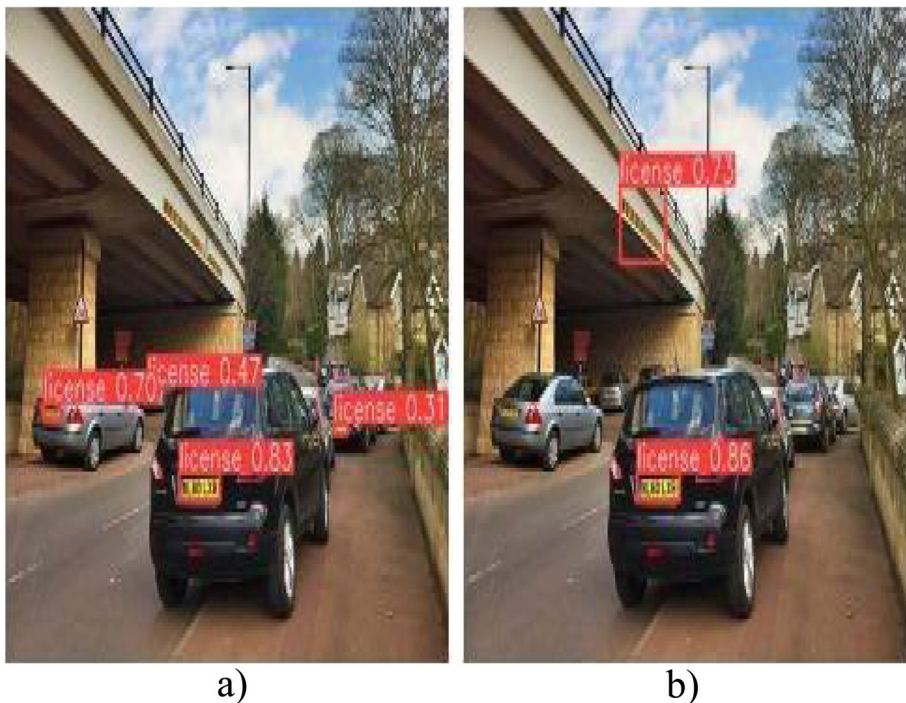


Fig. 9 a LPs detected by the 2Augmentation Model. b LPs detected with 1 Augmentation Model



Fig. 10 a LPs detected with 2 Augmentation Model. b LPs detected with 1 Augmentation Model



Fig. 11 Ideal scenario (back facing car) (a) 2 Augmentation model; b 1 Augmentation confidence about the detection of LP

predicted bounding box is considered a correct detection if it overlaps with the ground truth bounding box by at least 50%. So, mAP at IoU threshold of 0.5 indicates the average precision across different object categories when considering detections that have an IoU

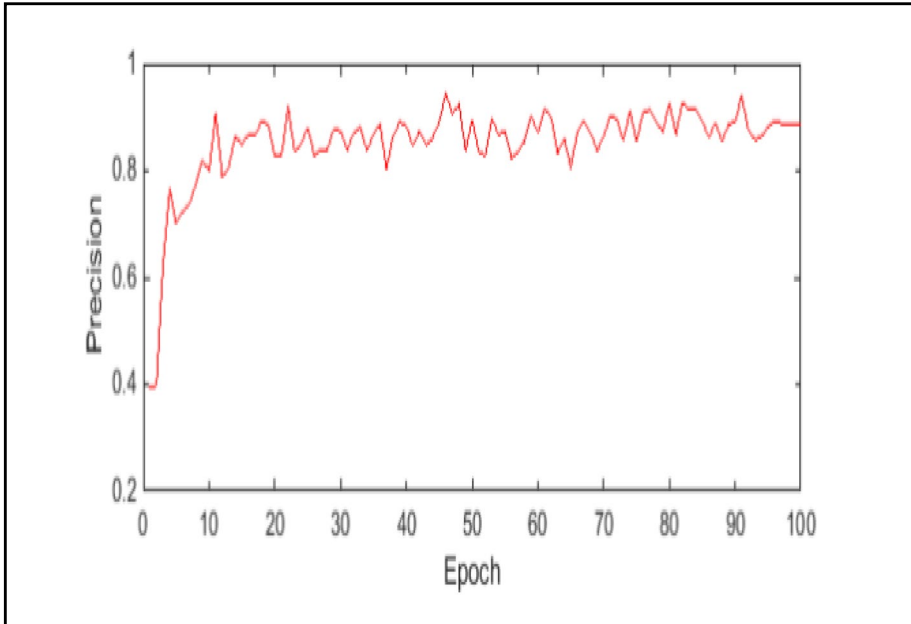


Fig. 12 Precision vs. Epoch Graph

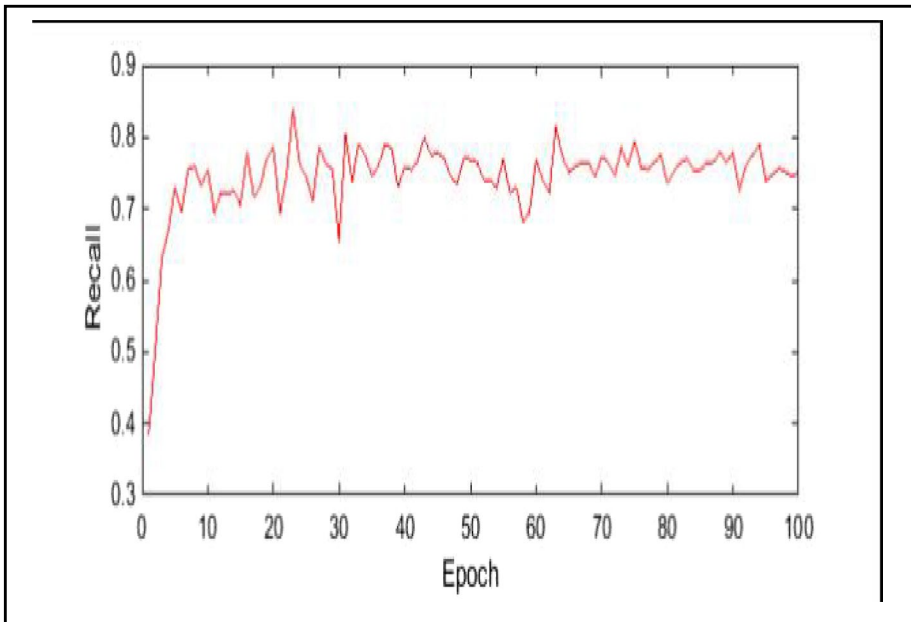


Fig. 13 Recall vs. Epoch Graph

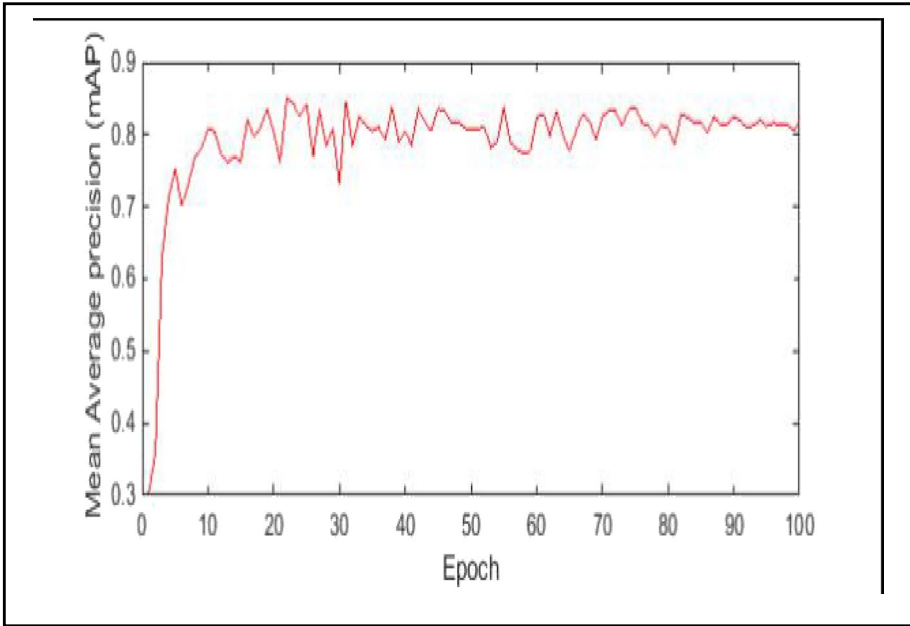


Fig. 14 Mean Average Precision vs. Epoch Graph

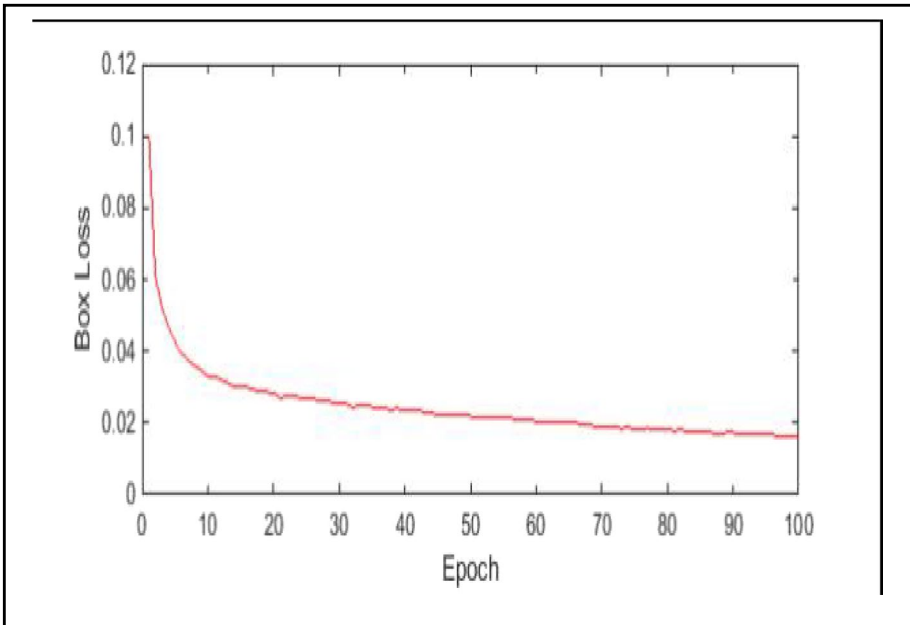


Fig. 15 Box Loss vs. Epoch Graph

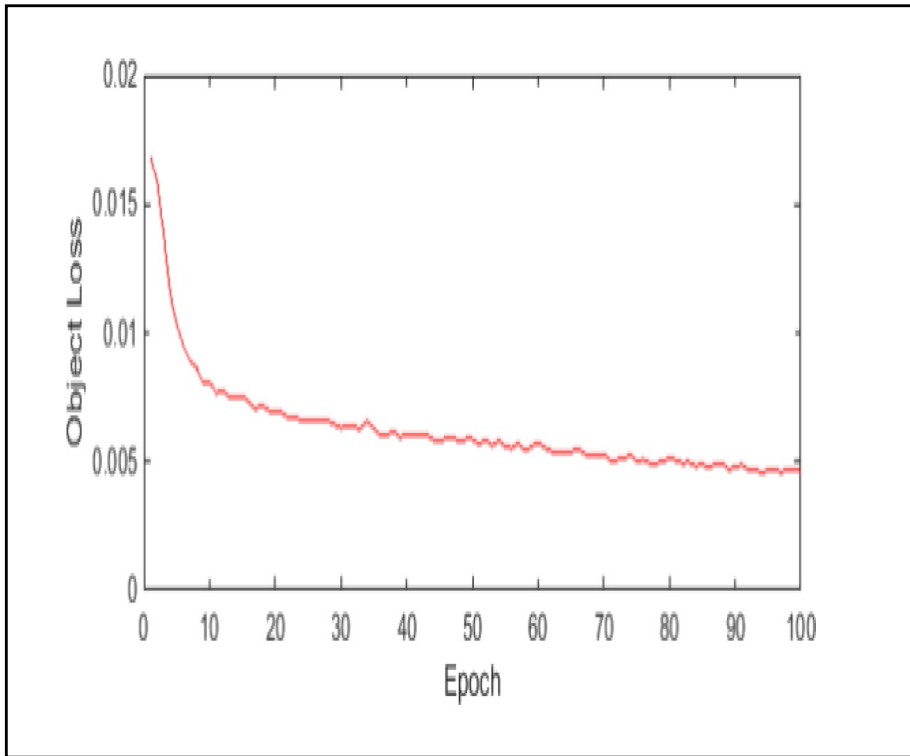


Fig. 16 Object Loss vs. Epoch Graph

Table 2 YOLO 5 performance measurement

Epoch	Precision	Recall	mAP@0.5
0	0.39565	0.38625	0.30356
15	0.86838	0.78	0.82081
30	0.83857	0.80517	0.84585
45	0.94487	0.77132	0.83361
60	0.91575	0.74	0.82892
75	0.90983	0.75675	0.81686
90	0.94022	0.725	0.81948
99	0.88895	0.75	0.81548

of at least 0.5 with the ground truth boxes. It provides a measure of the model's accuracy in detecting objects with a reasonable overlap with the ground truth. As noted from Table 2, the model attains favourable results at 45th epoch.

The object loss in YOLO encompasses three key components: Boundary box regression loss: which penalises inaccuracies in predicting box coordinates. Confidence of object presence penalising in correct confidence scores for object presence within boundary box. Classification loss: penalizing errors in predicting object categories. These components

collectively optimize object localization, confidence estimation and classification accuracy during training for robust object detection.

Figure 12 graph displays the flow of **Precision** over the course of model training. It can be observed from the graph that the precision of model improves **exponentially** during the starting 20 epochs of training. Then during the next 60 epochs we can see very little improvement in the metric. Then during the last 20 epochs, the metric stabilises a bit. Thus, more training epochs would not have resulted in any significant change(growth) in the metric. As observed from Fig. 12, the precision of model shows significant improvement upto 20 epoch and maximum precision is attained at 45th epoch. The graph in Fig. 13 displays the rate of change of **Recall** over the course of model training. It can be observed from the graph that the metric **increases exponentially** during the starting 45 epochs of training. Then the metric starts to take a detour and decreases in value over the next 15 epochs. This is observed due to the **presence of adverse kinds of images** which causes the model to **build better generalisation rules**. When the model has built special rules for all kinds of images, we can observe that the graph starts to rise again though we cannot see any major jump as the **growth stabilises** for the remaining epochs. Thus, more training epochs would not have resulted in any significant change(growth) in the metric.

The graph shown in Fig. 14 represents the flow of **Mean Average Precision** over the course of model training. It shows the same trend as in Precisions and Recall graphs. The bounding box is predicted with a loss function that gives the error between the predicted and ground truth bounding box. We can observe from the graph in Fig. 15 that the box loss metric makes a similar logistic function graph. At the very end, the loss metric starts to stabilize a bit. This is evident that more training would not have resulted in better metrics.

Object Loss is the confidence of object presence is the objectless loss (Binary Cross Entropy). The object loss metric graph behaves similar to box loss metric graph as shown in Figs. 15 and 16 with the difference being that object loss starts to stabilise way earlier than box loss. YOLOv5 Algorithm is designed in such a way that it automatically always picks the weights that maximises all 3 major metrics in object detection namely - **Precision, Recall and Mean Average Precision (mAP)**. After training the model, the model has been tested on various images clicked with different cameras angles and confidence level achieved for LP detection is more than 85%. Some of the results of tested images have been presented Fig. 17.

For increasing the robustness in the model, the images and videos from the dataset has been taken in different weather conditions and tested after adding 5% noise in images. It has been observed from Fig. 18 that more than 85% confidence level has been achieved.

5 Conclusion

In the present manuscript, we studied a model for detecting and recognising licence plate number under different conditions with different colours of number plates. The proposed model was based upon YOLO5 and proposed algorithm runs mainly into three phases: Vehicle detection, Licence Plate detection and recognition. To increase the robustness of model the dataset has been taken from different sources and noise has been added to images and then the model is trained. As observed, the detection results were further improved using two step augmentation. Precision, Recall, Mean Average Precision, box loss and object loss have been calculated and it has been observed that results are improving exponentially till 100 epochs and then stabilizes. Noise has been added into images



Fig. 17 Confidence level at different angles



Fig. 18 Confidence level after adding noise in model

so confidence level more than 85% has been achieved. At 45th epoch the model attains 0.83361 mAP@0.5 with precision of 0.94487 and recall of 0.77132. In subsequent research, the model can be improved by optimizing the backbone structure and other parameters [22].

Funding There is no funding source.

Data availability The dataset is publicly available Kaggle.

Declarations

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Pan SJ, Yang Q (2009) A survey on transfer learning. *IEEE Trans Knowl Data Eng* 22(10):1345–1359
2. Patel C, Shah D, Patel A (2013) Automatic number plate recognition system (ANPR): a survey. *Int J Comput Appl* 69(9):21–33. <https://doi.org/10.5120/11871-7665>
3. Anagnostopoulos C-NE et al (2008) License plate recognition from still images and video sequences: a survey. *IEEE Trans Intell Transp Syst* 9(3):377–391
4. Gou C et al (2015) Vehicle license plate recognition based on extremal regions and restricted boltzmann machines. *IEEE Trans Intell Transp Syst* 17(4):1096–1107
5. Hsu G-S, Chen J-C, Chung Y-Z (2012) Application-oriented license plate recognition. *IEEE Trans Veh Technol* 62(2):552–561
6. Ashtari AH, Nordin MJ, Fathy M (2014) An Iranian license plate recognition system based on color features. *IEEE Trans Intell Transp Syst* 15(4):1690–1705
7. Sarfraz MS et al (2013) Real-time automatic license plate recognition for CCTV forensic applications. *J Real-Time Image Proc* 8(3):285–295
8. Panahi R, Gholampour I (2016) Accurate detection and recognition of dirty vehicle plate numbers for high-speed applications. *IEEE Trans Intell Transp Syst* 18(4):767–779
9. Yuan Y et al (2016) A robust and efficient approach to license plate detection. *IEEE Trans Image Process* 26(3):1102–1114
10. Montazzolli S, Jung C (2017) Real-time brazilian license plate detection and recognition using deep convolutional neural networks. In: 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). IEEE
11. Azam S, Islam MM (2016) Automatic license plate detection in hazardous condition. *J Vis Commun Image Represent* 36:172–186
12. Rafique MA, Pedrycz W, Jeon M (2018) Vehicle license plate detection using region-based convolutional neural networks. *Soft Comput* 22(19):6429–6440
13. Svoboda P et al (2016) CNN for license plate motion deblurring. In: 2016 IEEE International Conference on Image Processing (ICIP). IEEE
14. Laroca R et al (2018) A robust real-time automatic license plate recognition based on the YOLO detector. In: 2018 International Joint Conference on Neural Networks (IJCNN). IEEE
15. Yang G et al (2020) Face mask recognition system with YOLOV5 based on image recognition. In: 2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE
16. <https://blog.roboflow.com/yolov5-improvements-and-evaluation/>. Accessed Dec 2022
17. Chen Z et al (2022) Plant disease recognition model based on improved YOLOv5. *Agronomy* 12(2):365
18. Zhu X et al (2021) TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In: Proceedings of the IEEE/CVF International Conference on Computer Vision

19. Available online at <https://www.kaggle.com/datasets/andrewmvd/car-plate-detection>. Accessed 5 Oct 2022
20. Available online at <https://www.kaggle.com/datasets/thamizhsterio/indian-license-plates>. Accessed 10 Oct 2022
21. Available on at <https://universe.roboflow.com/samrat-sahoo/license-plates-f8vsn>. Accessed 15 Oct 2022
22. Shashirangana J, Padmasiri H, Meedeniya D, Perera C (2021) Automated license plate recognition: a survey on methods and techniques. *IEEE Access* 9:11203–11225. <https://doi.org/10.1109/ACCESS.2020.3047929>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.