Available online at www.jrors.ir

Journal Of Radar and Optical Remote Sensing

JRORS 4 (2021) 30–40

License Plate Detection and Recognition based on Neural Networks in Complex Environments

N.Ameena Bibi[a], Purru Supriya [b*]

a. Assistant Professor(ECE),Government College of Technology,Coimbatore.

bApplied Electronics, Government College of Technology, Anna University, Coimbatore, India

Received 28 June 2021; Revised 24 August 2021; Accepted 13 November 2021

Abstract

Now a days due to the rapid advancement of economy around the world the count of vehicles increases day by day. Increase in the number of vehicles causes violation detection, road congestion, accidents at different traffic situations, uneven illumination, lighting and weather conditions. To overcome this issue license plate number is recognized but due to variations in license plate layout, font size of characters, tilted number plates, weather conditions, dirt plate and motion blur license plate recognition becomes difficult. License plate recognition has two main tasks, one is to detect the license plate and the other is to identify the license plate characters. By using region of interest license plate is detected. For recognition first tilted images are corrected using affine transformation and to improve the quality of a low-resolution image super resolution CNN is employed and connected component analysis, horizontal and vertical projection profile area used for separating each individuals' characters. Each individual character image is fed to the Convolutional Neural Network (CNN) for character extraction and for classification and the license plate is recognized using convolutional neural networks. The main aim of this paper is to recognize different plate layout with different conditions with minimum data set and less processing time with maximum efficiency.

Keywords*:* License Plate Recognition, Region of Interest, Horizontal and Vertical Projection, Convolutional Neural Network

1. Introduction

With the rapid development of economy around the world, the number of vehicles is incessantly greater than before, the traffic conditions are due to population growth and human needs, the use of vehicles has worsening, due to this many countries are faced with traffic congestion, frequent accidents and car theft which get giant pressure to the world. To overcome this license plate number is necessary. Car counting on highways, traffic violations detection, traffic monitoring, stolen car detection, journey

* Corresponding author Tel:
Email address: supriyapooty@gmail.com.

time measurement, travel time estimation, auto relying on roadways, petty criminal offences detection, and surveillance applications are just a few of the applications that use vehicle plate detection and recognition. In Intelligent Transport System, a vehicle's number plate is its only reliable identify. Intelligent transportation is a real-time, accurate, and efficient transportation management system.

For license plate Character recognition convolutional neural network is used. Other methods that are used for character recognition are deep learning methods, Optical character recognition, Maximization algorithm and wavelet transform, template matching, Adaptive Boosting and support vector machine.

In deep learning large number of datasets are required and it takes more time for processing. In Maximization algorithm selecting the correct number of decomposition level is difficult. In SVM and adaptive boosting Processing time is very high. In OCR the quality of the image can be lost during processing the image.

CNNs are able to learn relevant features from an image at different levels such as human brain. It automatically detects the important features without any human supervision. Compared to other recognition algorithms. The CNN layers are responsible for detecting low-level features such as edges and colors in the beginning layers and the layers at the end are responsible for finding out high-level features such as shapes. It will take only the important features by reducing the dimension of the image using maxpooling layer. CNN has very high accuracy in image recognition problems. By using CNN, a smaller number of datasets are required for recognition. Due to the smaller number of datasets Processing time is also very low.

2. License Plate Recognition Technology

In the field of image processing, license plate identification is a hotspot of research. Different algorithms have been proposed for finding the license plate number. The main disadvantage in these algorithms is huge number of datasets are required and processing time is high. To overcome this convolutional neural network is used for license plate recognition.

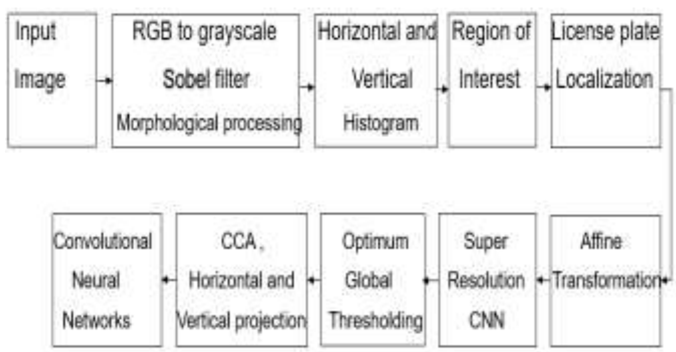


Fig 1. Block diagram of License Plate Recognition using convolutional neural networks

2.1. Preprocessing

RGB to Gray Scale: A video surveillance-based security system's ability to recognize license plates is critical. First Image is captured from a camera. Image is taken from fixed angle parallel to horizon. Input images are converted into greyscale image. Gray image is a single layer image from 0-255 where RGB is a three-layer image. Gray image is an 8bit image whereas RGB image is a 24bit image. Gray scale image is easy to process than RGB image by making a weighted sum of the red, green and blue components. RGB image is placed one above the other. Grayscale image needs 33%less memory than RGB image Grayscale images are significantly easier to work with in a variety of situations, such as morphological operations and image segmentation issues. In many morphological operation and image segmentation problem, it is easier to work with single layered image than a three-layered image. Grayscale means that it signifies only the intensity data of the light. Only the darkest black to the brightest white are displayed in grayscale images. In other words, the image contains only black, white, and grey colors, in which grey has multiple levels.

Sobel Filter: Sobel filter is used for edge detection and also to detect the orientations in an image. Sobel filter detect the edges in both horizontal(x) and vertical(y) direction. Gaussian filter plays a vital role in the entire process. In gaussian filter center having more weight than the rest We are scanning from top to bottom in the Y direction, and it will only detect horizontal edges in the image. Sobel X direction will detect the edges that will be processed from left to right. Multiply x and y with gaussian filter to detect Sobel x(Gx) and sobel y(Gy). Convolute GX and GY over the input image to calculate the value for one pixel at a time. Sobel all edges are selected by adding up the horizontal and the vertical edges.

Adding 45 degrees and 135 degrees with the horizontal and vertical direction of the sobel filter. Now, the horizontal and vertical direction produces new weight. This can be used to improve the performance of sobel edge detection. Because of the approximate gradient calculation, the Sobel operator is straightforward. It's employed in computer vision and image processing.

Morphological Processing: After edge detection there are a lot of unwanted edges which need to be eliminated and some of the edges which have gaps which need to be closed. Morphological processing is used to define the structure. Dilation will increase the white pixel of an image and reduce the black pixel. Holes are reduced in dilation.

Erosion will reduce the white pixel and rise the black pixels. Holes are raised in erosion. Adding or removing pixels from an object totally depends on the size and shape of the structuring element. If there occurs an overlapping then the pixels under the center site of the structuring element will be turned to 1 or black. If there is no overlapping then the structuring element will be turned to 0 or white. The structuring element is positioned at all possible sites in the image. For further processing morphological is needed(Weihong and Jiaoyang,2020).

2.2. Histogram Equalization

Histograms are used to increase the contrast in an image. Horizontal histogram and vertical histogram represent the column wise and row wise histogram. The sum of the differences in values between adjacent pixels in an image is represented by a histogram. For horizontal histogram, algorithm traverses through each column It starts with the second pixel from the top. The difference between the second and the first pixel is calculated If the difference exceeds the certain threshold, it is added to the total sum. It will move downward to calculate the difference between the third and the second pixels. It will move until the end of the column and calculates the total sum of differences between neighboring pixels. At the end an array containing column wise sum is created. The same process is followed to find the vertical correction rows.

Passing Histogram through Low Pass Filter

To avoid loss of data in forthcoming stages, excessive variations of histogram are flattened out. To flat out histogram is passed through low pass filter. It will detect only the extreme values in image. Areas having less histogram values are detached from the image. Because the area with the license plate has a plain background with alphanumeric characters. As a result, the disparity between neighboring pixels will be relatively large, especially around the margins of characters and numberplates. Therefore, horizontal and vertical histograms of lower values are not required anymore. As a result, they are filtered out by applying a dynamic threshold to both the horizontal and vertical histograms. The filtered histograms consist of the regions having the highest probability of containing the number plate.

Probable Candidate for Number Plate

In these regions, the one with the maximum histogram value is considered as the most probable candidate for number plate. The license plate boundaries of characters and number plate will have high histogram value(Huang et al., 2021).

2.3. Region of Interest

A region of interest is a subset of an image or a dataset identified for a particular purpose A recommended region from the original image is called a region of interest. The region of interest defines the bounds of an object under consideration in computer vision and optical character recognition. The concept of ROI is commonly used in many application areas. The entire area is processed with row wise and column wise to find a common region having maximum horizontal and vertical histogram value. In the entire area one having the highest histogram estimate will be considered as number plate by using region of interest. Mostly the number plate only will be having the highest histogram value compared to other images in the car(Wang et al.,2019).

2.4. Affine Transformation

Affine transformation is used to correct the tilted images that happen due to nonideal camera angles. All parallel lines, points, straight lines in the original image will still be same in output image. Angle between the lines and distance between the points will not be same in the output image. It maps variable (x1, y1) into new variable (x2, y2) by applying linear combinations of translation, scaling, rotation operations. The intensity value will be as same as the input image for output image. Affine transformation will operate only on the pixel values not on the intensity values.

Image Translation

Image translation is used to shift the license plate position from one place to another place.

$$J(x, y) = I(x+tx+y+ty) \quad (1)$$

tx=Translation of x location; ty=Translation of y location

Image Scaling

To increase or decrease the size image scaling is used in affine transformation. To perform scaling each coordinate is multiplied by a scaling factor.

$$J(x, y) = I(CXXx, CyXy) \quad (2)$$

Cx and Cy are the scaling factors to increase and decrease the size using x and y location.

Image Rotation

Image rotation takes an object and rotates it about a given axis, through some angle. With respect to x and y axis rotation of license plate is done. Based on the Euclidean distance image rotation is done(Arafat et al., 2018).

2.5. Super Resolution CNN

To improve the quality of low-resolution image super resolution technique is used. The end-to-end mapping is learned by a deep convolutional neural network. Bicubic interpolation is used to up sample low resolution to high resolution with the appropriate size. Then, deep CNNs are used to rebuild high-quality images.

The operations that make up the SRCNN are as follows:

i. Preprocessing: Up-scales Low resolution image to desired High resolution size.

Feature extraction: From the up-scaled LR image, a series of feature maps is extracted. Feature extraction produces n1 feature maps of low-resolution image

ii. Non-linear mapping: Nonlinear mapping uses convolutional layer with activation function. In this n1 feature map of low-resolution map is passed to convolutional layer with activation function to produce n2 feature map of high-resolution image.

iii. Reconstruction: Reconstruction layer adds both the n1 feature map and n2 feature map of to produce a high-resolution image.

Bicubic interpolation:

Bicubic interpolation is used to resize the image. It considers of 4x4 neighbourhood of known pixels. It produces smoother edges at the corners. It takes an input image and increases its width and height while preserving the image's quality.

$S_0(x)=a+bx+cx^2+dx^3$

a, b, c, d values are unknown. Bicubic interpolation are used to find these values and enlarge the image.

VGG16CNN:

In super resolution pretrained vgg16CNN is used to calculate the perceptual loss, peak signal to noise ratio and structural similarity index. Vgg16 has 16 layers so many features can be extracted. It has same filter size, stride and padding for all the convolutional layers and max pooling layers.

Performance Measures:

MSE: To measure the resemblance between two images, mean square error is used.

PSNR: It is used to measure quality of a lossy transformation.

Perceptual loss: Perceptual loss is computed by comparing two images based on high-level representations from pre-trained CNN model. The function is used to compare high-level differences between images, such as content and style variances.

SSIM: It's a way for determining how similar two images are in terms of contrast, structure, and luminance(Hilario et al.,2018).

2.6. OTSU Thresholding

For further segmentation process license plate image is converted into binary image. Otsu thresholding is used for converting greyscale image into binary image. The value of the threshold in Otsu thresholding is decided automatically rather than being chosen manually. By finding histogram of image threshold value is chosen. If saturation is greater image pixels are converted into white if saturation is smaller image pixels are converted into black.

2.7. Connected Component Analysis

It is used to find the similar objects in an image. This operation takes grey image as input. It detects the large sized connected foreground region. It figures out which parts of an image are connected physically. Pixels with same values belong to same object. CCA are clusters of pixels with same value, which are connected to each other. For further processing like horizontal and vertical projection profile CCA is used. In 4 connectivity pixels are connected if their edges touch horizontally and vertically(Shashirangana, et al.,2021).

2.8. Horizontal and Vertical Projection

The projection profile of an image along the horizontal axis is known as the horizontal projection profile. Horizontal projection counts the no of foreground pixels along the columns of the image. The text in a line with white pixels is represented by columns, which correspond to greater peaks in the histogram. The background image without text in a line having black pixels corresponds to lower peaks in histogram. columns that correspond to lower peaks in histogram can be used as a segmenting line to separate the character.

Vertical Projection Profile is the projection profile of an image along vertical axis. Rows represent the text in a line having high number of white pixels which corresponds to higher peaks in histogram. Rows represent the gap in-between the lines have high number of black pixels which corresponds to lower peaks in the histogram. Rows that correspond to lower peaks in histogram can be selected as segmenting lines to separate the characters(Shashirangana, et al.,2021).

2.9. Convolutional Neural Network

CNN is a Deep, feed forward artificial neural network that has been applied to analyse visual Imagery and it is a multilayered neural network that has been used to detect complicated features in data. Convolutional layer, RELU layer, Pooling layer, Dropout layer and fully connected layer are the layers that makes up CNN. CNN are more often utilized for feature extraction and classification.

A neural network is a collection of interconnected artificial neurons that communicate with one another. For pattern recognition, CNN is utilised. One or more convolutional layers, usually with a subsampling layer, are followed by one or more fully connected layers in a CNN. The weight combination extracts one colour while blurring the undesired noise. As the stride value is increased, the image size shrinks.

1. Convolutional Layer

First layer of CNN is Convolutional Layer. Extracts different features of input. First Convolutional layer extracts low level features like edges, lines and corners. Higher convolutional layer extracts high level features. Starting from top left corner of input each kernel moved from left to right, one element at a time. Once top right corner is reached, the kernel is moved one element in a downward direction until bottom right corner is reached. Filters are made to slide over the input image where filter is sometimes called neuron. Filters is also an array of number, the numbers are called weights or parameters. The feature map will be determined by the number of filters used.

2. Relu Layer

RELU is the Rectified Linear Units. Most deep learning network use rectified linear units for hidden layer. A rectified linear unit has the output 0 if the input is less than zero. I f the input is greater than zero, the output is equal to input.

$$F(x)=\max(x,0) \tag{3}$$

RELU layer leaves the size of volume unchanged.

3. Pooling Layer

To reduce the dimension of an image pooling layer is used. It reduces the resolution of features. It helps to reduce the parameters to great extent. When photos get too huge, the number of trainable parameters must be reduced. Pooling is done for the purpose of reducing spatial size of image. Pooling is done independently on each depth dimension

Two types of pooling,

a. Maximum pooling: Maximum value of four value is selected.

b. Average pooling: Average of four values in the region are selected.

In Max pooling the largest element from rectified feature map is considered. Most common form of pooling generally applied is Max Pooling.

The two-dimensional matrix joins pixels in both the horizontal and vertical directions, and the smaller dimensions help to minimise the parameters significantly. Depth, stride, and Zero Padding are the three parameters that influence the output volume size

i. Depth: It corresponds to number of filters. The raw image is sent into the first convolutional layer, which subsequently activates distinct neurons along the depth dimension.

ii. Stride: Stride will determine the pixels to be moved. When the stride is 1, move the filter one pixel at a time. When the stride is 2, filter is moved two pixels at a time. As we slide them around this produces smaller output as spatial volume.

iii. Zero Padding: Pad input volume with zero around the border. It allows to control the output volume's spatial size

$$\text{Zero padding } P=(F-1)/2 \tag{4}$$

When the stride is one, make sure the input and output volumes have the same size spatially.

4. Dropout Layer

When all of the features are connected to the FC layer, the training dataset may become overfit. After passing a dropout of 0.4, 40% of the nodes in the neural network are dropped out at random.

5. Fully Connected Layer

It is the Final layer of CNN. In the output layer, the Fully Connected layer employs the softmax activation function. The softmax function produces the output between 0 and1. It divides each output in half such that the total output is one. All elements of all features of the previous layer get used in the calculation of each element of each output feature. This layer sums a weight of the previous layer of features to determine the specific results. CNN are shift invariant. Same weight configuration is used across space. In CNN, same coefficients are used across multiple regions in space(Liu et al.,2019).

$$([W-F+2P]/S)+1 \tag{5}$$

W=input volume size
F= size of filter
P= number of padding applied
S= number of strides

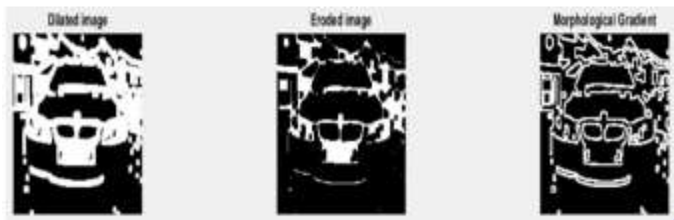
3. Experimental Results

The simulation of the Project is done using Python for different plate layout and is used to recognize the license plate using convolutional neural networks and different results are obtained for recognition of license plate.

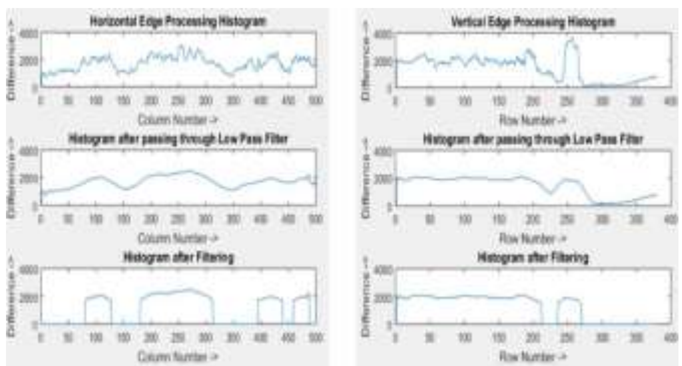
(a)

(b)

(c)



(d)

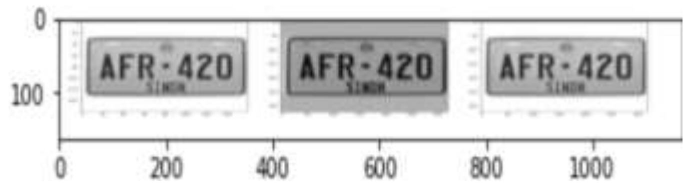


(e)

(f) (g)

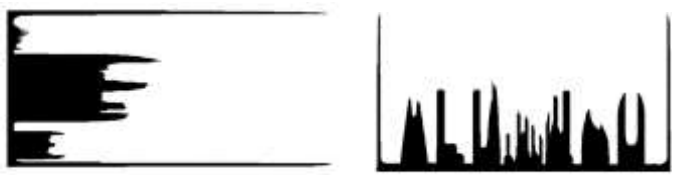
Figure 2. (a)input image (b) grey scale image (c) sobel filter (d) Morphological Processing (e) Horizontal and vertical edge processing (f) Probable candidate for number plate (g) license plate detection using region of interest

(g) (h)



(i)

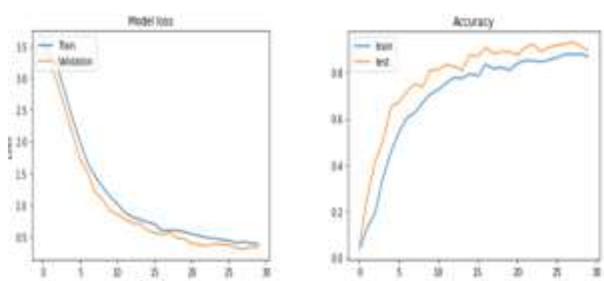
(j)



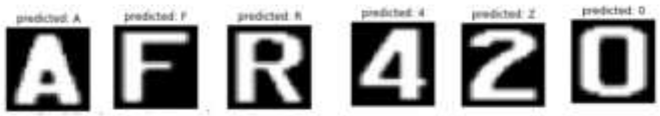
(k)

(l)

Figure 3. (h) affine transformation (i) Super resolution cnn (j) Binary image (k) Horizontal and vertical projection analysis (l) Character separation using projection analysis



(m)



(n)

Figure 4. (m)(n) license plate recognition using CNN

3.1. Accuracy for License Plate Recognition

Accuracy for license plate recognition is found out using confusion matrix
Accuracy= (10+2)/14=85.71%

Figure 5. Results of various images

3.2. Accuracy of Various Images

ORIGINAL IMAGE	PREDICTED IMAGE	ACCURACY
AFR20	AFRZ0	85.71%
DL4CAG9557	0L4CAG9557	90.9%
WBO2W688	WB02W888	90%
UP114BN4001	UP14BN4001	90.9%
SK049165	SK04916S	88.67%
IIPUL4	11PUL4	100%

4. Conclusion

License plate recognition is an important application in many areas. The License Plate Layout, font size of characters, different traffic situations, tilted car number plates, weather conditions, dirt license plate and motion blur due to fast moving vehicle due to this the recognition becomes difficult during many situations.

To overcome this convolutional neural network is used to recognize license plate. The license plate recognition system has two main tasks, one is to detect the license plate and the other is to identify the license plate characters. Localization of license plate is done using region of interest.

The license plate is recognized using convolutional neural networks with minimum number of datasets for both training and testing with less processing time and achieved accuracy of 92%.

References

Arafat, M. Y., Khairuddin, A. S. M., & Paramesran, R. (2018). A vehicular license plate recognition framework for skewed images. *KSII Transactions on Internet and Information Systems (TIIS), 12*(11), 5522-5540.

Chen, S. L., Yang, C., Ma, J. W., Chen, F., & Yin, X. C. (2019). Simultaneous end-to-end vehicle and license plate detection with multi-branch attention neural network. *IEEE Transactions on Intelligent Transportation systems*, *21*(9), 3686-3695.

Henry, C., Ahn, S. Y., & Lee, S. W. (2020). Multinational License Plate Recognition Using Generalized Character Sequence Detection. *IEEE Access*, *8*, 35185-35199.

Hilario Seibel Siome Golenstein, Andeerson, R. (2018). Super Resolution and License plate Recognition in Low -Quality surveillance Videos. *IEEE Access*, *5*, 20020-20035.

Huang, Q., Cai, Z., & Lan, T. (2021). A Single Neural Network for Mixed Style License Plate Detection and Recognition. *IEEE Access*, *9*, 21777-21785.

Liu, Y., Huang, H., Cao, J., & Huang, T. (2019). Convolutional neural networks-based intelligent recognition of Chinese license plates. *Soft Computing*, *22*(7), 2403-2419.

Li, H., Wang, P., & Shen, C. (2019). Toward end-to-end car license plate detection and recognition with deep neural networks. *IEEE Transactions on Intelligent Transportation systems*, *20*(3), 1126-1136.

Pustokhina, I. V., Pustokhin, D. A., Rodrigues, J. J., Gupta, D., Khanna, A., Shankar, K., ... & Joshi, G. P. (2020). Automatic vehicle license plate recognition using optimal K-means with convolutional neural network for intelligent transportation systems. *IEEE Access*, *8*, 92907-92917.

Shashirangana, J., Padmasiri, H., Meedeniya, D., & Perera, C. (2021). Automated license plate recognition: a survey on methods and techniques. *IEEE Access*, *9*, 11203-11225.

Wang, W., Yang, J., Chen, M., & Wang, P. (2019). A light CNN for end-to-end car license plates detection and recognition. *IEEE Access*, *7*, 173875-173883.

Weihong, W., & Jiaoyang, T. (2020). Research on license plate recognition algorithms based on deep learning in complex environment. *IEEE Access*, *8*, 91661-91675.