

Automatic License Plate Recognition for Distorted Images Using SRGAN



Anita Baral, Anupama Koirala, Sanjay Pantha, Rewant Pokhrel,
and Bishnu Hari Paudel

Abstract The quality of an image is always a limitation for automatic license plate recognition (ALPR) for its practical and accurate functioning. Distorted low-resolution (LR) images are one of the common issues encountered in ALPR systems. Most of the available algorithms in ALPR systems are effective under controlled circumstances or with intricate image capture systems. Authorities are looking for an improved ALPR system that is primarily concerned with dealing with distorted LR images. In this research, an improved ALPR system is proposed to generate high-resolution (HR) images from LR images using the super-resolution generative adversarial network (SRGAN). On a self-created Nepali dataset, both the networks (ALPR and SRGAN) are trained and assessed. Traditionally, for the identification of vehicles, images of license plates are localized, characters are segmented, and recognized. The accuracy of license plate recognition is heavily reliant on segmentation and recognition accuracy. ALPR with SRGAN comprises a super-resolution process in between. Super-resolution aids to improve the performance of the system by recreating detailed variation of the localized image, thus providing better end-to-end recognition.

Keywords Detection · Localization · Segmentation · Super-resolution · Generative adversarial network · Automatic license plate recognition

A. Baral (✉) · A. Koirala · S. Pantha · R. Pokhrel · B. H. Paudel
Institute of Engineering, Tribhuvan University, Pashchimanchal Campus, Pokhara, Nepal
e-mail: baralanita2@gmail.com

A. Koirala
e-mail: anukoira55@gmail.com

S. Pantha
e-mail: sanjaypantha.np@gmail.com

R. Pokhrel
e-mail: pokhrelrewant@gmail.com

B. H. Paudel
e-mail: bishnuhari@wrc.edu.n

1 Introduction

Automatic license plate recognition (ALPR) is an image processing technology that distinguishes vehicles by analyzing the images of their license plates. License plates are rectangular objects attached to a vehicle that display an official set of characters that identify the vehicle. Government organizations implement ALPR technologies for systematic law enforcement and easier investigations when required. ALPR technology allows the creation of a database of relevant data which speeds up time-consuming and labor-intensive manual differentiation of vehicle license plates against records of wanted, stolen, and other vehicles of interest.

ALPR algorithms work best in controlled environments or with complex image capture devices. In an uncontrolled environment, reading license plates accurately remains a significant challenge [1]. It is because a properly detailed image is required to provide enough information during segmentation and recognition to yield accurate results, which is not available in low-resolution (LR) images. In this research, a model is proposed which combines the ALPR system and super-resolution (SR) technique for LR distorted license plates to improve the license plate identification. SR is the task of estimating a high-resolution (HR) image from its LR counterpart [2]. SR can be achieved in many ways and algorithms, including the usage of deep convolutional neural networks and generative adversarial networks [3]. Super-resolution generative adversarial network (SRGAN) is used for super-resolution in this research. It provides a high-frequency detailed and perceptually satisfying counterpart to low-resolution license plates [2].

2 Related Work

Various research has been conducted in the automatic license plate recognition field and ways to improve its accuracy and effectiveness. Li et al. [1] proposed an automatic license plate (LP) detection and recognition system in natural background images using deep neural networks. The network can detect and recognize the plates in a single forward pass which avoids an error during transition and quicken plate processing. Pant et al. [4] proposed an ALPR for Nepali number plates with the use of support vector machines. Silva et al. [5] introduced automatic LP detection and recognition in unregulated conditions for multiple oblique and distorted LPs in a single image. The authors have also presented the solution of detecting LPs from different regions through the employment of manual annotations. In Balamurugan et al. [6], the authors exploit spline interpolation as a super-resolution technique in automatic number plate recognition system that detects plates from surveillance feed and up-scales the image using the SR method then implements OCR for recognition. Various techniques are being designed for the super-resolution of images. Ledig et al. [2] formulated photo-realistic single image super-resolution using generative adversarial network. A perceptual loss function consisting of content and adversarial

loss was proposed by the research to obtain photo-realistic images for the $4\times$ up-scaling factor. It has been seen that the implementation of SR techniques has been able to increase the performance of digital image processing techniques [7].

The substantial contribution of this work is a localized national ALPR system. We approach this problem through SRGAN-based techniques for LP detection and some optical character recognition (OCR)-based methods that can handle LP recognition.

3 System Architecture

3.1 Traditional ALPR System

Traditional automatic license plate recognition enacts three major steps: license plate detection and localization, character segmentation, and character recognition [8, 9]. Plate detection entails detecting the plate within the input image frame. Localization extracts the image of the detected plate from the input image and passes it through character segmentation, which further isolates each character. The isolated characters are now sent to OCR, which recognizes them (Fig. 1).

3.2 ALPR System for Distorted Images

The automatic license plate recognition system for distorted LR images enhances the efficiency and accuracy of the traditional ALPR system by refining the quality and resolution of the localized images, unlike the traditional system where the localized plate was directly fed to OCR after segmentation (Fig. 2).

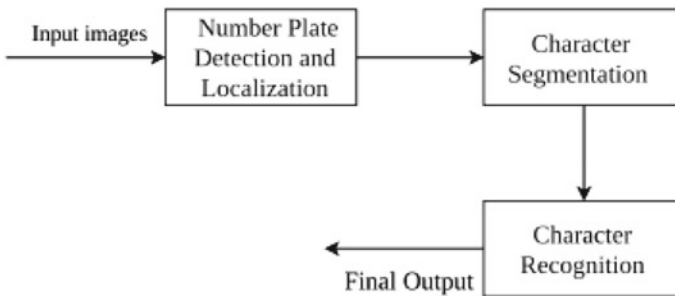


Fig. 1 Architecture of traditional ALPR system

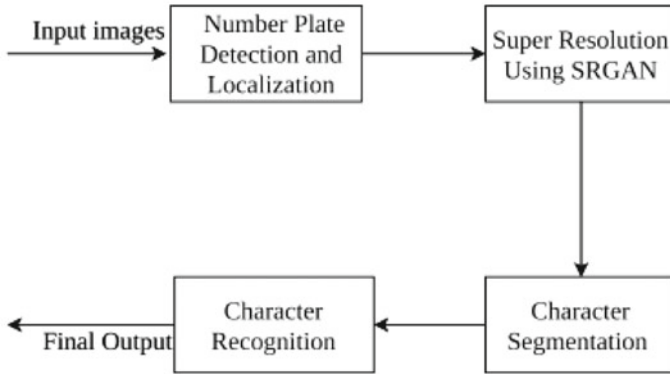


Fig. 2 Architecture of ALPR System for Distorted Images

4 Methodology

The raw input image is fed to the system and goes through various steps of preprocessing and processes to finally extract the characters. Starting from normalization and plate detection via Warped Planar Object Detection Network (WPOD-NET), a localized license plate image is obtained and fed to the SRGAN model. After some extra steps of grayscale and binary conversion, character segmentation is performed. Finally, segmented characters are fed to OCR model to obtain the result.

4.1 License Plate Detection

To detect license plates, a novel convolutional neural network (CNN) known as the Warped Planar Object Detection Network (WPOD-NET) [5] has been implemented. The WPOD-NET was obtained by incorporating ideas from You Only Look Once (YOLO), Single Shot MultiBox Detector (SSD), and Spatial Transformer Networks (STN) [5]. This network pursues to detect license plates in various distortions [5]. The following steps are involved for detecting and localizing the license plate:

Normalization In this step, we change the range of pixel intensity for the input image. The input image is converted in a range of pixel values between 0 and 1.

Plate Detection WPOD-NET is used to obtain coordinates of plate boundaries on the normalized image. The bounding box is drawn around the license plate with the help of the acquired coordinates. The coordinates are then used to extract the license plates from the images.

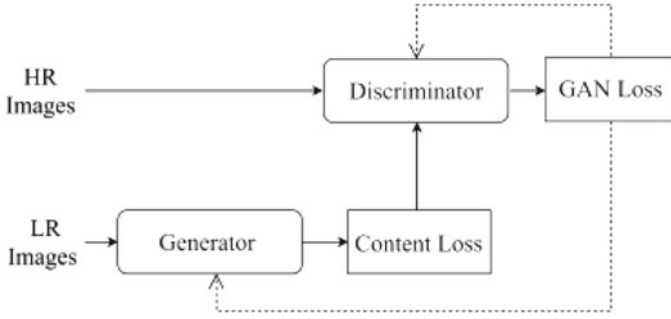


Fig. 3 Block diagram of SRGAN [10]

4.2 Super Resolution

SRGAN consists of two competing models, generator and discriminator [10], which can capture, copy and recreate detail variation within a dataset. The generator portion of SRGAN learns to create HR images by upscaling LR images through the incorporation of feedback from the discriminator output. The discriminator in an SRGAN is simply a classifier. It tries to distinguish real HR images from the HR images created by the generator. The discriminator comes across two types of HR images, real and fake. The real ones are the original HR images, whereas the fake ones are the upscaled HR images by the generator. SRGAN uses perceptual loss that consists of an adversarial and content loss. Perceptual loss function helps to obtain photo-realistic natural images. SRGAN relinquishes the use of pixel-wise mean square error (MSE) loss as content loss which is used by many other state-of-the-art systems because it cannot recover high-frequency contents from LR images which results in over smooth images for obtaining a high peak signal noise ratio (PSNR) (Fig. 3).

Perceptual Loss The overall loss in the generator network of SRGAN is based on perceptual loss l^{SR} , which is the weighted sum of the content loss l_x^{SR} and adversarial loss $l_{\text{Gen}}^{\text{SR}}$ [2].

$$l^{\text{SR}} = l_x^{\text{SR}} + 10^{-3} l_{\text{Gen}}^{\text{SR}} \quad (1)$$

Content Loss The content loss function for this SRGAN is defined as a VGG loss function $l_{\text{VGG}/i,j}^{\text{SR}}$, that is obtained from the pre-trained VGG network as depicted in Simonyan and Zisserman [11]. VGG loss is calculated as the Euclidean distance between feature map of an SR image $G_{\theta_G}(I^{\text{LR}})$ and the HR image I^{HR} [11].

$$l_{\text{VGG}/i,j}^{\text{SR}} = \frac{1}{W_{ij} H_{ij}} \sum_{x=1}^{W_{ij}} \sum_{y=1}^{H_{ij}} (\phi_{ij}(I^{\text{HR}})_{x,y} - \phi_{ij}(G_{\theta_G}(I^{\text{LR}}))_{x,y})^2 \quad (2)$$

Adversarial Loss Along with content loss, we add the generative component $l_{\text{Gen}}^{\text{SR}}$ of the GAN to the perceptual loss. By attempting to deceive the discriminator network,

the system is encouraged to prefer alternatives that reside on the manifold of natural images [2].

$$J_{\text{Gen}}^{\text{SR}} = \sum_{n=1}^N -\log D_{\theta_D}(G_{\theta_G}(I^{\text{LR}})) \quad (3)$$

4.3 Character Segmentation

Character segmentation segregates an image into constituted parts, each of which contains a character and can be excerpted for further processing. In this system, a contour-based approach has been implemented. This method is a gradient-based segmentation method that finds the boundaries based on the high gradient magnitudes. Before applying character segmentation, we have applied the following processing techniques to reduce noise and emphasize the key features of license characters.

Conversion to grayscale images To reduce the code complexity, for faster processing, and to ease the analysis of localized super-resolved license plates without keeping into consideration the color contrast, the images are primarily converted into grayscale.

Conversion to binary images The grayscale super-resolved license plate is then converted into an image with only black and white pixels. As a fundamental decree, this can be accomplished by establishing a threshold. If the pixel value exceeds the threshold, it is converted into a white pixel; otherwise, it is altered into a black pixel. Since the lighting condition is not uniform in the image, determining the threshold value becomes a crucial task.

4.4 Character Recognition

Individual characters from the localized plate are fed into the OCR system for prediction after being segmented. The OCR system employs learning and prediction algorithms based on support vector machines (SVM) to classify the segmented characters [12]. SVM is a versatile supervised machine learning algorithm that can perform linear or nonlinear classification, regression, and even outlier detection. The SVM algorithm's target is to determine a hyper-plane in an N-dimensional space that classifies the data points. The coordinates of each observation are used to calculate the support vectors. SVMs can yield precise and reliable classification results [12]. In this research, SVM is utilized to classify segmented characters among the twelve trained classes (०-९, ५ and ५).

5 Training Details

Datasets

- License Plate Dataset: There are approximately 1400 license plate images captured from various vehicles at different orientations and lighting conditions. The license plates are captured in the Gandaki zone, some in the underground parking while others in the outdoor parking.
- Character Dataset: There are a total of 6481 samples in the character dataset for twelve classes. The dataset is created by segmenting the characters using the contour-based approach in the localized license plate dataset and then manually categorizing them into twelve different classes.

Table 1 shows the number of representatives collected for each class.

SRGAN Training SRGAN model is trained approximately on 1100 localized number plates. The plates were all resized to a size of 200 * 280 for uniformity among the dataset. Two models were developed, one has an up-sampling factor of four while another has an up-scaling factor of eight. In SRGAN, perceptual quality is considered as a metric. The model has been trained for 1000 epochs since it generated a convincing result. Figure 4 illustrates the output of the SRGAN model for up-sampling factor four and eight.

OCR Training To train 4860 (75% of the total data-set) characters from twelve different classes, the Support Vector Classifier with polynomial kernel is used.

Table 1 Character dataset

Character	Sample
०	322
१	866
२	431
३	396
४	446
५	412
६	425
७	498
८	439
९	450
ग	873
घ	923
Total	6481

6 Experiments and Results

The purpose of this study is to assess the use of SRGAN to enhance the performance of the ALPR system. Figure 5 illustrates the overall working of the system. Our research is based on Nepali number plates of the Gandaki zone. Our system initially trains all the localized number plates in SRGAN, and then, the trained SRGAN model is used in conjunction with the traditional ALPR system (Fig. 5).

Segmentation Results The proposed system was evaluated on forty arbitrarily selected samples for license plate segmentation where the plates of the vehicles were not clear. Among the samples, twenty were highly pixelated and were up-scaled by factor eight in SRGAN, while the other twenty were up-scaled by factor four in SRGAN. Among the twenty images up-scaled by factor eight, seventeen samples were correctly segmented by the system, and among the images up-scaled by factor four, nineteen samples were correctly segmented.



Fig. 4 Output of SRGAN model for up-sampling factor four and eight

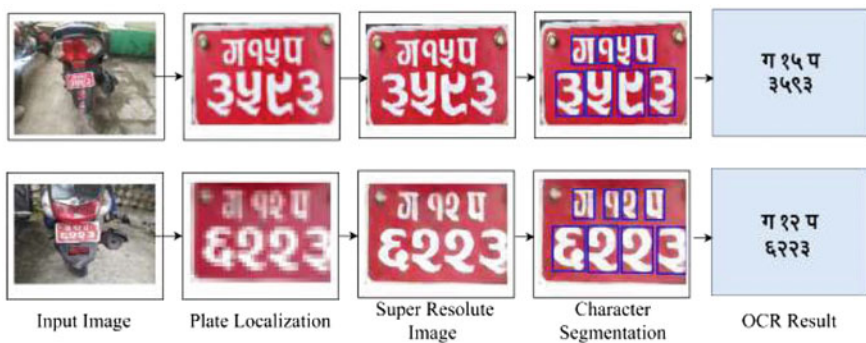


Fig. 5 Overall working of the system

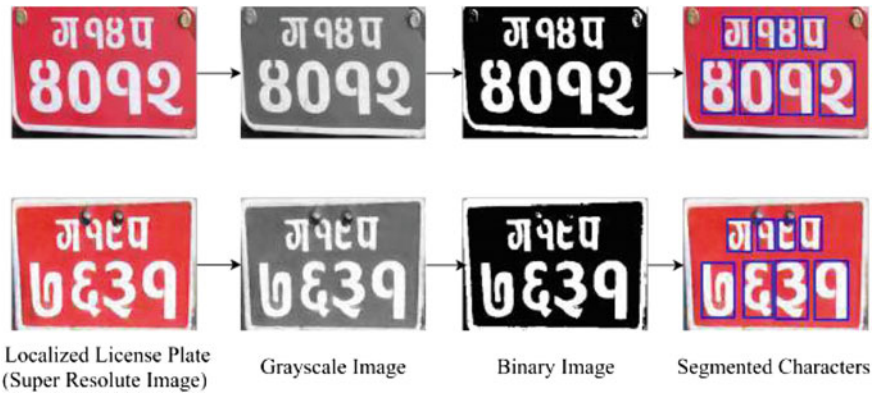


Fig. 6 Character segmentation

Table 2 Character recognition results

Metrics	Score
Accuracy	0.9858
Precision	0.9854
Recall	0.9842
<i>F1</i> -score	0.9847

Figure 6 illustrates the character segmentation for Super Resolute license plates. *OCR Result* Seventy-five percent of the data was used for training, while the remaining twenty-five percent was used for testing. Data samples were chosen at random and exclusively for training and testing in each experiment.

Table 2 shows the accuracy, precision, recall, and *F*-score of the OCR system. Table 3 displays the experiment's confusion matrix (CM).

Number Plate Recognition Results In total, forty images are tested. Twenty of the images were slightly distorted, while the other 20 were severely distorted. An experiment was conducted first on distorted images and then on up-scaled images. Former experiment shows an accuracy of 50% in slightly distorted images while in the case of heavily distorted images, the segmentation algorithm is not being able to detect the contours. On the later experiment that uses SRGAN with traditional ALPR, the evaluation result shows an accuracy of 90% in slightly distorted images that are up-scaled by factor four in SRGAN, and accuracy of 75% in heavily distorted images that are up-scaled by factor eight in SRGAN.

Table 3 Confusion matrix

Class	०	१	२	३	४	५	६	७	८	९	०	१	Total
०	81	0	0	1	2	0	0	0	0	0	0	1	85
१	0	209	0	0	0	0	0	0	0	0	0	0	209
२	0	0	101	0	0	0	0	0	0	0	0	0	101
३	1	1	0	96	0	0	0	0	0	0	0	0	98
४	0	0	0	0	108	0	0	0	1	0	1	0	110
५	0	2	0	0	0	101	0	0	0	0	0	0	103
६	0	0	0	0	0	0	114	0	0	0	0	0	114
७	1	1	0	0	0	0	0	138	0	0	0	0	140
८	0	1	0	0	0	0	0	0	90	0	0	0	91
९	1	1	1	0	0	1	0	0	0	120	0	0	124
०	1	1	0	1	0	0	0	0	0	0	216	2	221
१	0	0	0	0	0	0	1	0	0	0	0	223	224
Total	85	216	102	98	110	102	115	138	91	120	217	226	1620

7 Conclusion

In the real-world scenario, it is not always possible to get a visually understandable license plate taken by the commodity cameras which results in lower accuracy of the traditional ALPR models. ALPR models when combined with super-resolution techniques, SRGAN being one of them, can deliver better results. The research is carried out in the sector where traditional ALPR is failing to perform as expected. In severely and slightly distorted images, the ALPR system for distorted images achieved accuracy rates of 75% and 90%, respectively. In the event of slightly distorted images, the model improves accuracy by 40% when compared to traditional ALPR system, whereas in the case of heavily distorted images where the traditional system is failing, the proposed model shows an accuracy of 75%. The accuracy of model particularly in the case of heavily distorted images can be enhanced with further research in near future. Although the model's accuracy is a significant improvement over traditional ALPR, it can still perform better with better segmentation tools and an SRGAN model trained on a larger dataset and more number of epochs. SRGAN, along with ALPR, is a complex image processing model. This model becomes more viable when there is a larger number of dataset, i.e., ten times more than what has been used in this study. Although the model is complex, it has a great economic advantage in long term over high-end hardware components.

In Nepal, license plates are inconsistent, discolored, and mud-splattered. The above-mentioned factors affect segmentation, recognition accuracy, and, ultimately, the overall accuracy of the system. This study is being conducted for the number plates of Nepal's Gandaki zone, the research can be expanded across the country. Along with

this, the dataset consists of only bike and scooter license plates; other automobiles should be included for the system to be realistic. Improving the aforementioned issues can result in a system with cutting-edge performance. Automatic license plate recognition for distorted images thus combines the components of traditional ALPR systems with SRGAN to achieve better results in real-world scenarios.

References

1. Li H, Wang P, Shen C (2017) Towards end-to-end car license plates detection and recognition with deep neural networks. [arXiv:1709.08828](https://arxiv.org/abs/1709.08828) [cs], Sep. 2017
2. Ledig C et al (2017) Photo-realistic single image super-resolution using a generative adversarial network. [arXiv:1609.04802](https://arxiv.org/abs/1609.04802) [cs, stat], May 2017
3. Dong C, Loy CC, He K, Tang X (2015) Image super-resolution using deep convolutional networks. [arXiv:1501.00092](https://arxiv.org/abs/1501.00092) [cs], Jul. 2015
4. Pant AK, Gyawali PK, Acharya S (2015) Automatic Nepali number plate recognition with support vector machines. In: Presented at the international conference on software, knowledge, information management and applications (SKIMA), Dec. 2015
5. Silva SM, Jung CR (2018) License plate detection and recognition in unconstrained scenarios. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds) Computer Vision ECCV 2018, vol 11216. Springer International Publishing, Cham, pp 593–609
6. Balamurugan G, Punniakodi S, Rajeswari K, Arulalan V (2015) Automatic number plate recognition system using super-resolution technique. In: 2015 International conference on computing and communications technologies (ICCCCT), Chennai, India, Feb. 2015, pp. 273–277. <https://doi.org/10.1109/ICCCCT2.2015.7292759>
7. Singh A, Singh J (2016) Super resolution applications in modern digital image processing. *IJCA* 150(2):6–8. <https://doi.org/10.5120/ijca2016911458>
8. Indira K, Mohan KV, Nikhilashwary T (2019) Automatic license plate recognition. In: Bhat-tacharyya S, Mukherjee A, Bhaumik H, Das S, Yoshida K (eds) Recent trends in signal and image processing, vol 727. Springer Singapore, Singapore, pp 67–77
9. Renuka devi D, Kanagapushpavalli D (2011) Automatic license plate recognition. In: 3rd International conference on trendz in information sciences computing (TISC2011), Dec. 2011, pp. 75–78. <https://doi.org/10.1109/TISC.2011.6169088>
10. Goodfellow IJ et al (2014) Generative adversarial networks. [arXiv:1406.2661](https://arxiv.org/abs/1406.2661) [cs, stat], Jun. 2014
11. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) [cs], Apr. 2015
12. Auria L, Moro RA (2008) Support vector machines (SVM) as a technique for solvency analysis. *SSRN J*. <https://doi.org/10.2139/ssrn.1424949>