XVII International Scientific Conference on Industrial Systems (IS'17)
Novi Sad, Serbia, October 4. – 6. 2017.
University of Novi Sad, Faculty of Technical Sciences,
Department for Industrial Engineering and Management
Available online at http://www.iim.ftn.uns.ac.rs/is17

# Deep Learning Driven Plates Recognition System

**Marko Arsenović**
(Teaching Assistant, University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia, arsenovic@uns.ac.rs)
**Srđan Sladojević**
(Assistant Professor, University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia, sladojevic@uns.ac.rs)
**Andraš Anderla**
(Assistant Professor, University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia, andras@uns.ac.rs)
**Darko Stefanović**
(Assistant Professor, University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia, darkoste@uns.ac.rs)

**Abstract**

*Optical character recognition (OCR) represents one of well-studied areas of computer vision and machine learning. Due to the complexity of the problem there is always room for improvement. Usually, there are specialized solutions for certain types of character recognition tasks in many areas of industry: banking, legal, government, education and etc. Authors of this paper propose a new approach of developing the deep learning based method for plate recognition system. Proposed method consists of combining the state-of-the-art deep neural networks for both of the essential tasks: plate detection and character recognition which were trained and validated on artificially generated vehicle plates' dataset. Complete development process is described in details and a test system is implemented, as well. The proposed system is tested in the real-time environment and achieved overall accuracy is 98% for plate detection, 97% for letters recognition and 94.5% for digit recognition. The proposed system has almost real-time response and it is suitable for image and video based recognitions.*

**Key words:** *ocr, plate recognition, CNN*

## 1. INTRODUCTION

Plenty of text-based data is embedded in images or videos in a real-time environment. In order to directly use this information placed on them by some software applications, text needs to be extracted automatically. Optical character recognition (OCR) represents the process of transforming the text information from images or videos in the text format. This computer vision task is challenging due to the great diversity in the text representation on the images and videos. Text information could be handwritten or in various fonts, noised or partially displayed. OCR based systems apply different approaches in order to response to these difficulties and to receive the text information with enough accuracy to be used in the natural scenes. Because of these practical demands, most of these systems are not general-purpose solutions, but rather specialized ones.

Document image analysis represents one of the areas where OCR techniques are highly applied. In many areas of industry there are many handwritten or machine printed documents that need to be digitalized. In order to reduce time, automatic solutions are vital. These OCR solutions vary in their complexity and constraints which could be quality of paper and ink, background, whether documents use fixed-font or multi-font and many others [1]. OCR systems for the document analysis are applied in banking where they automatically process checks by scanning them and making all of the financial transferring without human involvement. Another usage of these systems is in legal industry where all the paper files are scanned and stored in the database ready for use in electronic format. Healthcare professionals deal with large amount of paper forms from patients and insurances. OCR based solutions extract all these forms' information and store it into databases making this data easily searchable and accessible. OCR powered processing solutions could be found in other industries which deal with similar problems such as education, government and etc.

Due to the continuous improvements in OCR methods and machine learning in general, OCR driven solutions are used for vehicles' plates recognition, house number recognition, extracting data from commercial hoardings, improving search algorithms on the web, for data scrapping and etc.

OCR based solution presented in this paper uses state-of-the-art deep learning methods for plate number

recognition process that could be used as a smart parking management system but it is also applicable in another system where this kind of data is needed.
The rest of the paper is organized as follows: Section 2 presents related work, Section 3 presents methods of developing proposed OCR system, Section 4 presents achieved results and discussion related to them, and Section 5 holds our conclusion.

## 2. RELATED WORK

Due to the fact that OCR systems represent one of the areas where computer vision and machine learning were applied the most, prior work include various approaches and methodologies.

Up to now, OCR tasks are solved by applying several steps which include text detection, segmentation along with different pre-processing and character recognition as the final step of the process which involves feature extractions and classification. Different text detection methods depend on feature engineering techniques that are using: boundary features (boundary of license plate for example), colour features (colour of the plate) [2], specific colour of the plate and texture features [3] (colour transitions on the plate) or character features [4].

Segmentation process could include detecting character contours, pixel connectivity or combination of these features [5, 6].

There are various methods of feature extractions for character recognition which depend on image representation forms (grey level, binary, vector), such as template matching and zoning [7]. There are modern approaches in character recognition which involve using the most advanced techniques in deep learning. As an example, deep convolutional neural networks (CNN) are used for multi-digit number recognition task [8] with high accuracy rate. Authors in the review paper [9] presented currently used techniques and procedures' steps for plate number detection tasks. These approaches include the methodologies mentioned in this section. Current advances in deep learning motivated the authors of the paper [10] to exploit CNN for license plate recognition task which resulted in achieving over 98% accuracy for digits and 96% for letters. Based on this information, authors of this paper use deep CNN for character recognition process.

Authors in [11] proposed parking allocation system which relies on RFID for car detection and OCR plate recognition. They used Sobel, horizontal and vertical edge detectors along with convolution matrices for OCR. Authors in [12] proposed OCR and RFID based smart parking management system. In the proposed system, OCR is developed by applying pre-processing and plate detection using image filtering, segmentation and template matching. Accuracy of both systems mentioned above were not specified.

Authors in [13] proposed intelligent speed bump system based on license plate recognition. Prior to appropriate image pre-processing and segmentation, they exploited Tesseract OCR tool [14] for character recognition task.

## 3. MATERIALS AND METHODS

The proposed deep learning based OCR plate detection system contains several components which could be used as a general-purpose solution. First component of the system is the end device which needs to have video camera and network module. Second component of the system is the OCR plate recognition system which will be described in detail further in this section. Third component of the proposed system is the server side, which holds the database and the business logic of the system. Proposed general system solution is displayed in Figure 1.

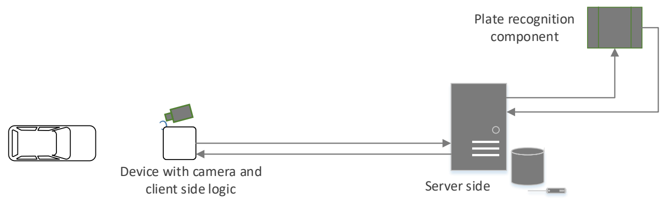


**Figure 1**. General system solution

### 3.1 Developing OCR component

Developing the OCR component for the proposed plate detection system consists of several important steps which includes preparing the dataset with appropriate image pre-processing, defining the plate object detector and character recognition as a last step.

#### 3.1.1 *Dataset and pre-processing*

Quality of the training and validation dataset is essential element of developing OCR component with high accuracy. For the purpose of this research artificially generated dataset was used. Process of generating the dataset consisted from gathering background images (around ten thousand of images), generating vehicle number plates and noising the image. For the background images, two publicly available databases were used, SUN database [15] and Stanford background database [16].

Plates were generated automatically by developed Python script that uses random text, plate colours, and fonts along with several predefined ratios for bounding box which are usually in the use for motors, trucks and cars. For these values, only European vehicle and traffic standards were used in this research, but this approach could be adopted for another standard also. In order to adopt the deep CNN for different angles and positions of the vehicle and also for different lighting conditions, certain pre-processing on the dataset's images was applied. Plate images were transformed by applying affine transformations based on the random scale, roll, pitch and translation. When the affine transformation was applied, position of the bounding box was stored in the file system in Pascal VOC xml format which is later used for training the CNN plate detector.

Last step of preparation of the image is noising which was applied in order to prevent the CNN to adopt on sharp edges but also to simulate the real-time environments' conditions: camera noise and lightning variations. Samples of the generated dataset are displayed in Figure 2.


**Figure 2**. Samples of the dataset

### 3.1.2 *CNN Plate detector*

First step of the proposed OCR system is plate detection on the image or video frame. For plate detection, standard approaches usually include feature engineered algorithms, as it is mentioned in the Section 2. In this research, specialized deep CNN was used due to its high accuracy but most significantly for its generalization power which plays the important role in real-time environments. Network's architecture, "You only look once" (YOLO), used for this object detection task is proposed by the authors in [17]. The main reason for choosing this approach for plate detection is its speed. This method is able to process images at 45 frames per second in real-time. Plate detector CNN in this paper is developed in Tensorflow [18], machine learning framework, trained, validated and fine-tuned on the dataset described in the previous section by using xml bounding boxes as the ground truth.

### *3.1.3 CNN Character recognition*

Plate recognition is the final step of the OCR component for the proposed system. Detected vehicle plate from the previous CNN presents the input of the plate recognition CNN. For this specific character recognition task, network architecture proposed in the paper [19] was used with certain modification to fit for plate character recognition for Serbian standard where the proposed system was tested. Output layer of the network used in this paper has 7·62 neurons, 7 characters represented by 62 neurons (digits, small and capital letters). Due to better compatibility of the entire system, plate recognition CNN was also implemented in Tensorflow.

Process of plate recognition is displayed in Figure 3.

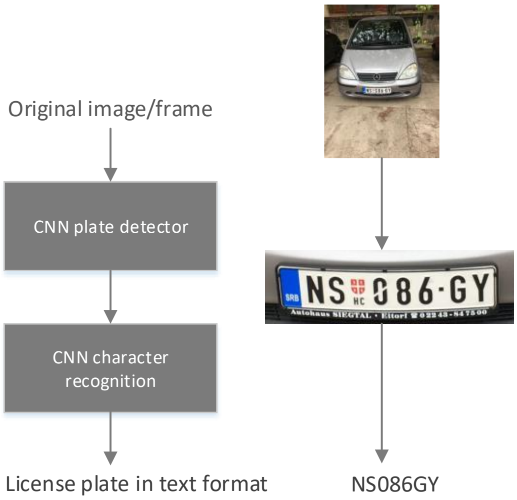


**Figure 3**. Plate recognition process

### *3.1.4 Plate recognition test system*

System used for testing the developed OCR component is the implementation of the general system solution mentioned at the beginning of this section. It uses Android mobile device as the client side. Image can be captured or uploaded from the gallery and sent to the server side which holds the OCR component. OCR component process the image through the CNNs which results in the plate's text which is sent as the response back to the Android. Figure 4. represents the client side application for testing the proposed system.

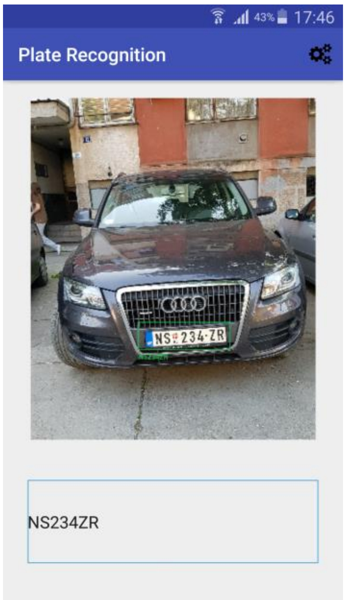


**Figure 4**. Mobile application

## 4. RESULTS AND DISCUSSION

OCR component's accuracy was assessed for the both essential steps: plate detection and character recognition. In the testing environment, 10-fold-cross validation was applied during the networks developing, final overall accuracy for the plate detector CNN was 98.5% and 96.8% for character recognition CNN.

Proposed system and OCR component was tested in the real-time environment using the test solution mentioned in the previous section. Fifty vehicles in different position and light conditions were captured by mobile phones and sent to the server which responded in a timely fashion. Overall accuracy for the plate detection was 98%, 97% for letter recognition and 94.5% for the digit recognition. Slight difference in fonts on the real plates and the plates from the dataset for several digits resulted in lower accuracy, Figure 5.

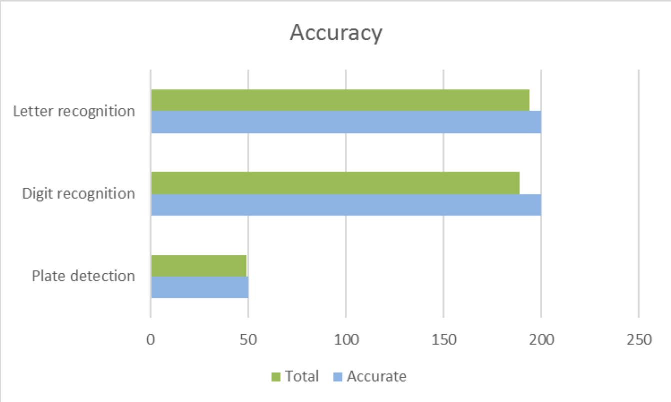


**Figure 5**. CNNs accuracy in real-time environments

## 5. CONCLUSION

Up to now there are many approaches used for OCR-based tasks. Due to the complexity of the character recognition problems which depend on various factors (paper or image quality, amount of noise, lightening conditions, fonts and etc.), solutions are usually specialized with constraints for certain the type of character recognition problems.

In this paper, deep learning based plate recognition system is proposed. The system consists of several important components which could be used as a general-purpose solution. Test system was developed in order to assess the OCR component's accuracy and response time. With high overall accuracy and almost instant server response, proposed system could be applied in the real-time environment. Limitations of the current work is the OCR component, which is designed for the certain type of vehicle plates – 7 characters' plates with several types of fonts which are used in the area where the initial test took place.

Future work will include developing the general solution for the plate detection by introducing new component for character recognition and modifications of the CNN-based character recognizer. Researching newer approaches in image pre-processing for simulating real-world situations could improve accuracy of the system. One of the unexplored aspects of this research is changing the system architecture in the way of porting the OCR component on the client side, whether it is mobile or some specialized embedded device. In that way, widespread usage could be achieved by excluding the server side component or at least provide the possibility of the offline mode work regime.

## 6. REFERENCES

[1] Baird, Henry S., and Karl Tombre. "The Evolution of Document Image Analysis." Handbook of Document Image Processing and Recognition. Springer London, 2014. 63-71.

[2] Yao, Cong, et al. "Rotation-invariant features for multi-oriented text detection in natural images." PloS one 8.8 (2013): e70173.

[3] K. Jung, K. I. Kim, and Anil. K. Jain, "Text information extraction in images and video: a survey", Pattern Recognition, vol. 37, 2004, pp.977-997. Sad

[4] Yao, Cong, et al. "Strokelets: A learned multi-scale representation for scene text recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.

[5] Haralick, Robert M., and Linda G. Shapiro. "Image segmentation techniques." Computer vision, graphics, and image processing 29.1 (1985): 100-132.

[6] Dave, Namrata. "Segmentation methods for handwritten character recognition." International journal of signal processing, image processing and pattern recognition 8.4 (2015): 155-164.

[7] Trier, Øivind Due, Anil K. Jain, and Torfinn Taxt. "Feature extraction methods for character recognition-a survey." Pattern recognition 29.4 (1996): 641-662.

[8] Goodfellow, Ian J., et al. "Multi-digit number recognition from street view imagery using deep convolutional neural networks." arXiv preprint arXiv:1312.6082 (2013).

[9] Du, Shan, et al. "Automatic license plate recognition (ALPR): A state-of-the-art review." IEEE Transactions on Circuits and Systems for Video Technology 23.2 (2013): 311-325.

[10] Menotti, David, et al. "Vehicle license plate recognition with random convolutional networks." Graphics, Patterns and Images (SIBGRAPI), 2014 27th SIBGRAPI Conference on. IEEE, 2014.

[11] Jondhale, Asmita, Gautami Das, and Samadhan Sonavane. "OCR and RFID enabled vehicle identification and parking allocation system." Pervasive Computing (ICPC), 2015 International Conference on. IEEE, 2015.

[12] Joshi, Yadnesh, et al. "Smart parking management system using RFID and OCR." Energy Systems and Applications, 2015 International Conference on. IEEE, 2015.

[13] Ho, Cheng-Yu, Huei-Yung Lin, and Lu-Ting Wu. "Intelligent speed bump system with dynamic license plate recognition." Industrial Technology (ICIT), 2016 IEEE International Conference on. IEEE, 2016.

[14] Smith, Ray. "An overview of the Tesseract OCR engine." Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on. Vol. 2. IEEE, 2007.

[15] Xiao, Jianxiong, et al. "Sun database: Large-scale scene recognition from abbey to zoo." Computer vision and pattern recognition (CVPR), 2010 IEEE conference on. IEEE, 2010.

[16] Heitz, Geremy, et al. "Cascaded classification models: Combining models for holistic scene understanding." Advances in Neural Information Processing Systems. 2009.

[17] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

[18] Abadi, Martín, et al. "Tensorflow: Large-scale machine learning on heterogeneous distributed systems." arXiv preprint arXiv:1603.04467 (2016).

[19] Stark, Fabian, et al. "Captcha recognition with active deep learning." Workshop New Challenges in Neural Computation 2015. 2015.