International Journal of Advanced Smart Convergence Vol.9 No.1 113-120 (2020)
http://dx.doi.org/10.7236/IJASC.2020.9.1.113

IJASC 20-1-14

Semi-Supervised Learning Based Anomaly Detection for License Plate OCR in Real Time Video

Bada Kim, Junyoung Heo

M.E., Department of Computer Engineering, Hansung University, Republic of Korea
Associate Professor, Department of Computer Engineering, Hansung University, Republic of Korea
qkek984@hansung.ac.kr, jyheo@hansung.ac.kr

Abstract

Recently, the license plate OCR system has been commercialized in a variety of fields and preferred utilizing low-cost embedded systems using only cameras. This system has a high recognition rate of about 98% or more for the environments such as parking lots where non-vehicle is restricted; however, the environments where non-vehicle objects are not restricted, the recognition rate is about 50% to 70%. This low performance is due to the changes in the environment by non-vehicle objects in real-time situations that occur anomaly data which is similar to the license plates. In this paper, we implement the appropriate anomaly detection based on semi-supervised learning for the license plate OCR system in the real-time environment where the appearance of non-vehicle objects is not restricted. In the experiment, we compare systems which anomaly detection is not implemented in the preceding research with the proposed system in this paper. As a result, the systems which anomaly detection is not implemented had a recognition rate of 77%; however, the systems with the semi-supervised learning based on anomaly detection had 88% of recognition rate. Using the techniques of anomaly detection based on the semi-supervised learning was effective in detecting anomaly data and it was helpful to improve the recognition rate of real-time situations.

Keywords: *Anomaly Detection, Deep Learning, License Plate Recognition, Semi-Supervised Learning, Optical Character Recognition (OCR)*

1. Introduction

Optical Character Recognition (OCR) is a technology which converts a letter written by a machine or person into a character that is recognizable by the computer. Recently, it has been evolving into a more accurate technology thanks to the development of deep learning [1]. It is commercialized in various fields such as reading printed materials, capturing credit cards when registering, and recognizing license plates [2]. For the license plate recognition, low-cost embedded systems with using only cameras are preferred without the usage of additional sensors. If a sensor is used to detect objects, its units cost much higher than only camera-based systems. On the other hand, only camera-based systems are inexpensive and can perform well with using deep

Manuscript Received: January. 21, 2020 / Revised: January. 27, 2020 / Accepted: January. 30, 2020
Corresponding Author: jyheo@hansung.ac.kr
Tel: +82-2-760-8039, Fax: +82-2-6919-2148
Associate Professor, Department of Computer Engineering, Hansung University, Republic of Korea

learning [3]. Low-cost embedded systems are implemented based on System on Chip (SoC) board. SoC board is a technology that can implement the self-integration system. These SoC board based embedded systems are as shown in Figure 1. It is a design that classifies the input of image data by the camera on the SoC board.

Figure 1. SoC board based embedded system architecture

The automatic parking management system is a typical example of a commercial license plate OCR system. It has a high recognition rate of about 98% in the environments where non-vehicle objects are restricted. However, the recognition rate is about 50% to 70% in the environments where non-vehicle objects are not restricted such as roads and alleys where environmental conditions are not constant [4]. This low performance is because anomaly data similar to license plate interferes with recognition in real-time. Even if pre-processing is done, there are a lot of different types of these anomaly data. Therefore, to implement license plate OCR in real time video, it has to be accompanied by high recognition rates and algorithms for anomaly detection [5].

Most OCR systems focus on class classification for input characters, and also general license plate OCR systems are the same. However, license plate OCR systems used in the environments where non-vehicle objects are not restricted may collect unnecessary noise images as well. If this noise becomes an input value, the Deep Learning network outputs to the most similar class. These outputs are wrong answers.

In this paper, we implement appropriate anomaly detection for the license plate OCR system in the environment where non-vehicle objects are not restricted. To detect anomaly that classifies input noise, semi-supervised learning is used to train both normal data and anomaly data. The experiment is conducted in two type. First of all, the verification accuracy of the supervised learning model and the semi-supervised learning model are compared to measure the performance of anomaly detection. To evaluate real-time recognition performance, the OCR performance of supervised learning based system and semi-supervised learning based system in real-time video are analyzed.

In this paper, we describe the preceding license plate OCR system and abnormal detection in Chapter 2. Chapter 3 describes the semi-supervised learning based anomaly detection, training, and predicted value analysis methods proposed in this paper. Chapter 4 presents the results of the experiment and the performance evaluation of the license plate OCR system proposed in this paper, and finally Chapter 5 concludes the paper by discussing conclusions and future research

2. Related Work

2.1 Preceding License Plate OCR

Most license plate OCR systems have a five-step process [6]. The first step is, the process of capturing the vehicle(Figure 2 (a)), in which a license plate text image is obtained at the front or rear of the vehicle. The next step is, the localization(Figure 2 (b)) process of license plates , which extracts the location of license plates with the characters in the images. The third step is the pre-processing (Figure 2 (c)) process in which license plate text is divided into individual segment and the slope of text is adjusted, then the shade and noise are removed. The fourth stage is the process of license plate recognition (Figure 2 (d)) that classifies the specific character of an individual segment. Finally, the license plate output (Figure 2 (e)) process combines recognized characters into a single string to output the entire of license plate text.

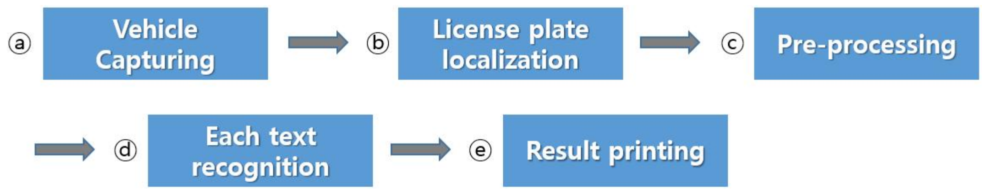

Figure 2. Process of license plate OCR

In the case of a preceding study [7], the license plate OCR system was implemented by supervised learning 43,221 the license plate image on the Convolutional Neural Networks (CNN) based. The experiment was conducted at non-real time with 97.57% accuracy. In this paper, we implement CNN based supervised learning license plate OCR system proposed by [7]. To compare anomaly detection performance in the real-time environments, supervised learning based systems are compared with this paper's proposed semi-supervised-based systems in Chapter 4 experiments.

2.2 Anomaly Detection

Anomaly detection means identifying data which is not normal in the process of data analysis [8]. Observations of normal data can be clustered and correlated, but anomaly data has quite different patterns of distribution away from the cluster. The basic principle of anomaly detection is to identify these anomaly points.

Anomaly detection is significant in prediction function of Deep Learning. The predictive function of general in-depth learning is to present a given input value as an output value, which was determined to be the most similar among previously trained labels. If anomaly data enters the input when there is no anomaly detection function, the deep learning model has no choice but to return the output value to a similar label of the class that already exists.

Anomaly detection can be implemented in several ways. There are three main ways of implementation: supervised learning, unsupervised learning, and mixed methods [9]. In this paper, we detect the anomaly data based on semi-supervised learning, which is the method using supervised learning is applied. In addition to the prediction model, the supervised learning based anomaly detection places additional models for detecting anomaly data, Whereas, semi-supervised learning based anomaly detection is a way to predict data by training both normal and anomaly data together in one model

3. Design of anomaly detection

3.1 Semi-Supervised Learning

General supervised learning is a technique that produces a model by training labelled data. For semi-supervised learning, both labeled data and unlabeled data are used as training data for one model [10]. Semi-supervised learning has trained all the data, it helped to distinguish the association among the embedded feature vectors more clearly, especially for the labeled data [11]. Thanks to these features, semi-supervised learning is often used to achieve high learning accuracy in relatively low data situations.

In this paper, the characteristic training all labels that can be candidates for input values is useful for detecting anomaly data, so semi-supervised learning is used for the purpose of implementing anomaly detection. This not only makes the labeled data in feature space more clearly distinguishable, but anomaly data of various shapes can also be classified as unlabeled data with features that have not been generalized. Labeled data trains normal data, which are segmented license plate text. Unlabeled data trains anomaly data and random noise that may be confused with license plate text.

3.2 Data training

For this paper, which is aimed at number plate OCR in real-time video, numerous anomaly data are accompanied by input values. Figure 3 is an example of the data of license plate area extracted from the localization process. Figure 3 (a) is a typical license plate data, but Figure 3 (b) is localized due to the features similar to the license plate. Figure 3 (b) has several types of phone numbers or company logos with text characteristics, as well as objects such as vehicle radiator grille with the features similar to text's edge changes. Therefore, the labelled data for semi-supervised learning consists of segmented data for each text in the Figure 3 (a) data. Similarly, unlabeled data consists of anomaly and random noise data of the type in which segmented text of Figure 3 (b).

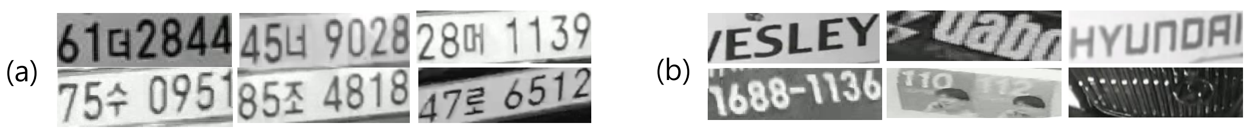

Figure 3. Localized data

The text of the license plate in Republic of Korea is composed of symbols by vehicle type, usage, and the names of cities [12]. Symbols by the vehicle types are numeric symbols consisting of 0 to 9, Symbols by usage use a total of 40 types of single character, and symbols by the names of cities use two-letters names totaling 17 types. As in Figure 4, license plates include non-business vehicles, normal business vehicles, and large business vehicles, and the position of each symbol is fixed according to the type of license plate. A non-business license plate consists of a combination of numeric and single character symbols, and a business license plate is additionally combined with a two-letter city name symbol. Types of business license plates are divided into normal and large vehicles. The normal business vehicle license plate has a vertical city name, while the large business vehicle license plate has a horizontal name of city. The license plate OCR system in this paper trains four models, separated by each symbol type of license plate text. Unlabeled data are trained together in all models for anomaly detection. Each trained model is used to suit the position of symbols to be predicted. Models separated by the symbol types are trained only on labels which can be output according to the input, thus preventing anomaly output by other symbol outputs can be avoided.

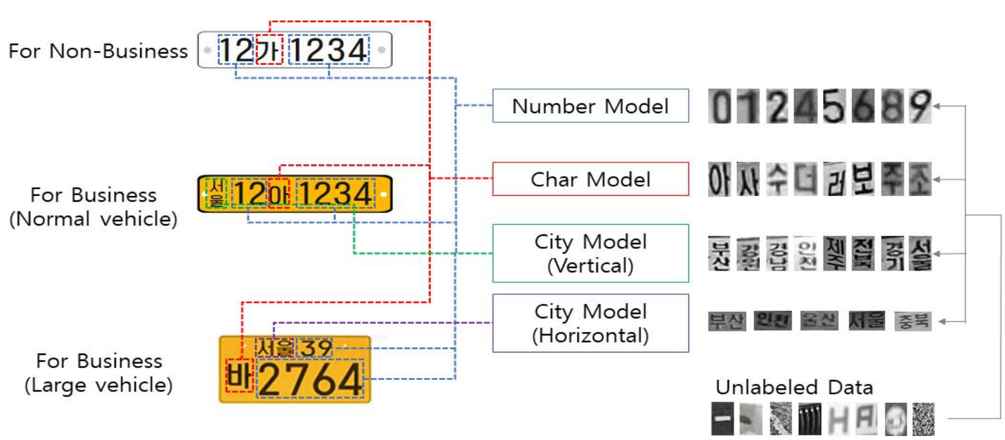

Figure 4. Each text type model of license plate

3.3 Analysis technique of predicted result

The real-time license plate OCR in this paper trains the model as in SqueezeNet v1.1 [13]. It is a lightweight design that operates on the SoC board. Most general network architectures use the Softmax function on the last layer to return the probability that the total sum is normalized to 1 [14]. In this paper, to output the return value to the logit score, a model with the last

layer is modified to Relu and then used, in predicting the data. This is because for numerical comparison of values with other frame results, relative comparable logit score is more suitable than normalized probability.

In this paper, we make a conclusion by comparatively analyzing the predicted logit score from the frames. This analysis can filter out anomaly data from some of the predicted results to make a more reliable full number plate text. The analysis method divides the groups of predicted results by the same location of the recognized license plate text in each frame. Then, the G-Score (Group Score) of the divided group is calculated as equation (1). The equation (1) adds up average score of a divided group and weights (W) proportional to the average score. Finally, the entire text is derived from combining the predicted results of the highest G-score for each digit area. G-score is a method that combines the average score and the weight of the scores in the group. The weights are higher for the groups with higher prediction frequency and the groups are recognized as higher scoring averages; in contrast, the weights are lower for the groups with lower prediction frequency and they are recognized as lower scoring averages.

$$G - Score(S_i) = \frac{\sum_{k=0}^{n_i} S_{ik}}{n_i} + \sum_{k=1}^{n_i} S_{ik} \cdot w \tag{1}$$

In equation (1), S_i is score set for each prediction group, W is the random weight

4. Experiment

4.1 Experiment Criteria

The model uses a training model and predicting model that is modified to suit this system based on the squeezeNet v1.1 model. Image channels are input/output in grayscale, loss functions use a categorical crossentropy suitable for multi-class classification, and optimization functions use adam. The code is written in python 3.6 and the deep learning framework operated on a keras 2.2.4 based.

The data set is shown in table 1. For the comparative experiment, the four types of models are trained as supervised learning based and semi-supervised learning methods respectively. The supervised learning based method trains only the training data, and the semi-supervised learning method trains the training data together with the anomaly data.

Table 1. Data set

Model	Training data	Training anomaly data	Verification data (A type)	Verification data (B type)
Vertical city	9421	576	972	1029
Horizontal city	9376	537	931	984
Number	24278	925	2366	2458
Char	41597	1227	4098	4220

The evaluation is conducted by two different ways. First, the model validation evaluation measures the accuracy of the supervised learning based and semi-supervised learning models in two types of verification sets. Verification data type A did not include anomaly data, and verification data type B included anomaly data. About 10% of the training anomaly data is included within the verification data type B. Second, the real-time recognition performance evaluation evaluates the implemented license plate OCR system. The experiment evaluates the integrated performance of the supervised learning based system and the proposed system of this paper. The evaluation data is a real-time video in which vehicles on real-world roads are captured. The video captured the images of 150 vehicles during the day and 181 during the night.

4.2 Validation Evaluation of Model

Table 2 shows the results of the model validation evaluation. As a result of the evaluation of Type A, which does not include anomaly data, supervised learning based trained model showed 99%, 97%, 97% and 96% accuracy, respectively, in the order of numbers, characters, vertical city names, and horizontal city name models. The semi-supervised learning based trained model showed 98%, 96%, 96%, and 95% of accuracy, respectively, in the same order as before. The two learning methods showed a difference of about 1 percent in the environment not including anomaly data.

Table 2. Verification test

Technique	Model	A type Accuracy	B type Accuracy
Supervised Learning Based	Number	0.9979	0.9479
	Char	0.9768	0.9358
	Vertical city	0.9712	0.9203
	Horizontal city	0.9667	0.9187
Semi-Supervised Learning Based	Number	0.9831	0.9805
	Char	0.9675	0.9630
	Vertical city	0.9619	0.9553
	Horizontal city	0.9549	0.9502

The results of the verification data B type evaluation that includes the anomaly data were slightly different. supervised learning based trained model showed 94%, 93%, 92% and 91% of accuracy, respectively, in the same order as before. The semi-supervised learning based trained model showed 98%, 96%, 95%, and 95% of accuracy, respectively, in the same order as before. Most supervised learning based trained models' performances have significantly reduced by 5%, While most semi-supervised learning based trained models have maintained their original performance. Considering that about 10% of the training anomaly data rate is included in the verification data type B, it is expected that the accuracy of supervised learning model will be decreased while the rate of the anomaly data increased. Overall, the supervised learning based models were not able to classify anomaly data, and the semi-supervised based models were found to be capable of classifying the anomaly data

4.3 Performance Evaluation in Real-Time Video

For real-time performance evaluation, precision, recall, F1-score, and accuracy are measured. The measurement criteria are the confusion matrix shown in Figure 5. If the license plate text recognition is correct, it is classified as True Positive (TP), but if it is incorrect or fails to localize, it is classified as False Negative (FN). If anomaly data is not detected, it is classified as false positives (FP). If it is detected, it is classified as true negative (TN). In summary, TP and TN correspond to the correct answers, positive data and negative data are correctly classified, FN and FP are incorrect, and positive data and negative data are not correctly classified.

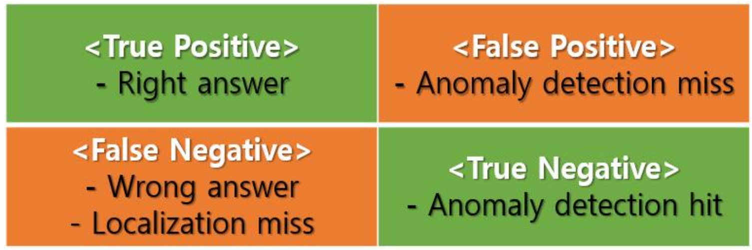

Figure 5. A confusion matrix of this paper

Analysis of real-time recognition performance is shown in Table 3. Supervised learning based showed 77% accuracy while the semi-supervised learning base showed 88% accuracy.

To verify the effect of detection on anomaly data, it is important to check the precision calculated by the TP and FP ratios. Supervised learning based was 90% precision, but the semi-supervised learning based detected most of the anomaly data, showing 97% precision. In this paper, Recall is not high on both sides, as noise or localization effects from real-time situations other than anomaly data are not taken into account.

Table 3. Performance test in real time video

System	Precision	Recall	F1-score	Accuracy
Supervised Learning Based	0.9014	0.8504	0.8752	0.7794
Semi-Supervised Learning Based	0.9726	0.9018	0.9356	0.8821

5. Conclusion

Anomaly data adversely affects OCR performance of license plate in real-time video. In the paper, semi-supervised learning-based license plate OCR system was implemented to detect anomaly data. For the comparative experiments, supervised learning based license plate OCR system were implemented in the preceding research, and the models were verified in the static situations without anomaly data, showing the excellent accuracy of about 96% to 99%. However, as a result of a recognition experiment in real-time video on the road, supervised learning based has been decreased in the accuracy of 77% due to the effects of anomaly data. Unlike the supervised learning based, the semi-supervised learning-based license plate OCR proposed in this paper showed the excellent performance of about 95 to 98% under any situations. In addition, the proposed system in this paper showed 88% accuracy in real-time video of roads. Given the precision of the proposed system in this paper was 97%, it was possible to classify most anomaly data, and it was found that semi-supervised based anomaly detection was effective.

The limitation of this paper are 90% of recall of the proposed system, which is relatively lower than precision. The reason for the low recall is that there is a lot of missing localization in real-time situations. The missing of localization has an effect on the reduction of FN. In real time, the changes in noise or lighting interfere with the proposal of localized areas. To improve accuracy, future research will address the functions affecting FN. If further studies are made on how to improve the performance of pre-processing and localization algorithms, it is expected that a more accurate license plate OCR system will be implemented.

Acknowledgement

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 19DPIW-C153746-01).

References

[1] Cho, Wooyeong, et al., "A Comparative Study on OCR using Super-Resolution for Small Fonts," *International journal of advanced smart convergence*, Vol. 8, No.3, pp. 95-101, Oct 2019.
DOI: https://doi.org/10.7236/IJASC.2019.8.3.95

[2] Zhong, Z., Jin, L., and Xie, Z., "High performance offline handwritten chinese character recognition using googlenet and directional feature maps," *International Conference on Document Analysis and Recognition (ICDAR)*, IEEE, pp.846-850, Aug 2015.
DOI: https://doi.org/10.1109/ICDAR.2015.7333881

[3] Lee, S., Son, K., Kim, H., and Park, J, "Car plate recognition based on CNN using embedded system with GPU," *In 2017 10th International Conference on Human System Interactions (HSI)*, pp. 239-241, Jul 2017.
DOI: https://doi.org/10.1109/HSI.2017.8005037

[4] ETRI, Purpose and Necessity of Technology Transfer. *https://itec.etri.re.kr/itec/sub02/sub02_01_1.do?t_id=7101-2018-01492.*

[5] Lopresti, D, "Optical character recognition errors and their effects on natural language processing," *International Journal on Document Analysis and Recognition (IJDAR),* Vol. 12, No. 3, pp. 141-151, Sep 2009.
DOI: https://doi.org/10.1007/s10032-009-0094-8

[6] Sonavane, Kiran, Badal Soni, and Umakanta Majhi. "Survey on Automatic Number Plate Recognition (ANR)," *International Journal of Computer Applications*, Vol. 125, No.6, Sep 2015.
DOI: https://doi.org/10.5120/ijca2015905920

[7] Jung-Hwan Kim and Joon-Hong Lim, "License Plate Detection and Recognition Algorithm using Deep Learning," *The Journal of Institute of Electrical and Electronic Engineers*, Vol. 23, No. 2, pp. 642-651, June 2019.
DOI: https://doi.org/10.7471/ikeee.2019.23.2.642

[8] Ahmed, Mohiuddin, Abdun Naser Mahmood, and Jiankun Hu. "A survey of network anomaly detection techniques," *Journal of Network and Computer Applications*, Vol. 60, pp. 19-31, Jan 2016.
DOI: https://doi.org/10.1016/j.jnca.2015.11.016

[9] Park, Chang-Mok. "Dam Sensor Outlier Detection using Mixed Prediction Model and Supervised Learning." *International journal of advanced smart convergence,* Vol. 7, No. 1, pp. 24-32, Mar 2018.
DOI: https://doi.org/10.7236/IJASC.2018.7.1.4

[10] Görnitz, Nico, et al. "Toward supervised anomaly detection," *Journal of Artificial Intelligence Research,* Vol. 46, pp. 235-262, Feb 2013.
DOI: https://doi.org/10.1613/jair.3623

[11] Haeusser, P., Mordvintsev, A., & Cremers, D, "Learning by Association—A Versatile Semi-Supervised Training Method for Neural Networks," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 89-98, Jul 2017.
DOI: https://doi.org/10.1109/CVPR.2017.74

[12] Ministry of Land, Infrastructure and Transport, Notice on the Standard of Registration Number Plate. *http://www.law.go.kr/admRulLsInfoP.do?admRulSeq=2100000176243.*

[13] Iandola, Forrest N., et al. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," *arXiv preprint arXiv:1602.07360*, Feb 2016.
arXiv: https://arxiv.org/abs/1602.07360

[14] Jang, Eric, Shixiang Gu, and Ben Poole. "Categorical reparameterization with gumbel-softmax," *arXiv preprint arXiv:1611.01144*, Nov 2016.
arXiv: https://arxiv.org/abs/1611.01144