Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 18-19, 2023

A Deep Learning Approach to Nameplate Detection and Text Recognition using YOLOv5 and Tesseract OCR

Seul Lee [a], Daeho Kim [a], Jaepil Ko [b*]
[a] *KEPCO E&C, 269, Hyeoksin-ro, Gimcheon-si, Gyeongsangbuk-do, Republic of Korea*
[b] *Kumoh National Institute of Technology, 61 Daehak-ro, Gumi, Gyeongbuk, 39177, Korea*
[*] *Corresponding author: nonezero@kumoh.ac.kr*

1. Introduction

All equipment of a nuclear power plant shall be affixed with a nameplate (i.e., a label) engraving the description and tag number of the equipment. The nameplate is also attached to the top of parts such as indicator and switches. The operator can identify the equipment or component by checking the tag number and the description of the nameplate.

A tag number is very important information for nuclear power plant design and operation. All information related to the equipment, such as device supplier information, installation location, and related drawings may be obtained through the tag number. The technology for recognizing the information engraved on the nameplate based on computer vision has various fields of application. For example, it can be implemented as a system that acquires tag numbers with a camera of the mobile device and immediately provides maintenance information and specifications of the equipment. It can also be applied in interface with augmented reality (AR) and cyber plant system.

Convolutional Neural Networks (CNNs) have become increasingly important in the field of object recognition, particularly for image-based tasks, because they perform breakthrough in tasks such as ImageNet classification tasks [1]. Object detection has evolved into two main paradigms: two-stage detectors and single-stage detectors. Single-stage detectors have gained increasing attention due to their ability to perform real-time object recognition with reduced computational complexity. Among the single-stage object detectors, YOLO (You Only Look Once) has emerged as an influential approach in the field of real-time object detection. YOLO's breakthrough design significantly reduces computational costs compared to conventional two-stage detectors by processing images with a single forward propagation through the network, enabling fast and accurate object detection [2].

Along with image-based object recognition, Optical Character Recognition (OCR) is another essential technology that has seen significant advancements in recent years. Tesseract OCR, an open-source OCR engine developed by Google, has gained widespread adoption due to its high accuracy and support for numerous languages [3]. The combination of object detection algorithms and OCR engines allows robust and versatile solutions in tasks of detecting and recognizing text of nameplate in complex scenes of nuclear power plant site setting.

This paper proposes a real-time method of detecting nameplate objects with YOLOv5 and extracting nameplate text by applying Tesseract OCR to the detected nameplate area. This methodology aims for fast operation and high accuracy to facilitate application in the field and in various other implementations.

2. Method and Experiment

The proposed method consists of two main stages: nameplate detection using YOLOv5 and text recognition using Tesseract OCR as shown in Fig. 1. The following subsections detail each stage.

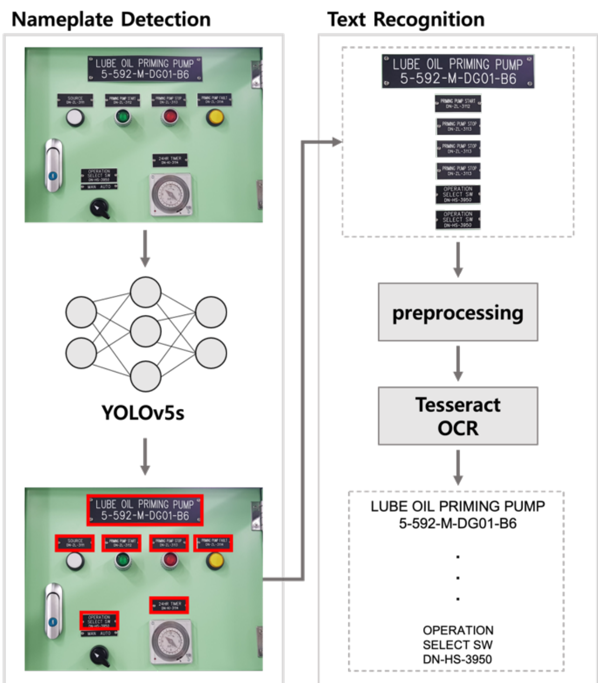

Fig.1. Framework of nameplate detection and text recognition

2.1 Nameplate Detection

It is a step of obtaining coordinates of bounding box by detecting the nameplate area from the local control panel image. To detect nameplate in the images, YOLOv5 algorithm was adopted. Among the YOLOv5 models, YOLOv5s model was trained, validated, and tested using image of local control panel.

2.1.1 Dataset. The dataset consisted of 398 image of local control panel containing different types of nameplates. Each image data included from 1 object (nameplate) to 55 objects. The images were annotated manually using labelImg [4]. The dataset was divided into training, validation, and test sets with a ratio of 60:20:20. Data augmentation such as flipping [5], and Mosaic augmentation [6] were applied to increase the diversity of the dataset and improve the generalization capabilities.

2.1.2 Training. The YOLOv5s is a lightweight variant of the YOLOv5 family, specifically designed for efficient and faster object detection while maintaining a good balance between speed and accuracy.
The architecture of YOLOv5s is based on the CSPNet (Cross Stage Hierarchical Networks) backbone [7] and PANet (Path Aggregation Network) for the neck [8]. It uses a series of convolutional layers, residual blocks, and skip connections to learn discriminative features from the input images.
The YOLOv5s model was trained on the nameplate dataset for 150 epochs, with 15 iterations per epoch, using a training image size of 1280x1280 pixels. The initial learning rate was set to 0.001 and Adam was adopted as optimizer.

2.1.3 Object Detection. After training, the nameplate is detected using the model and best weights. The input images were resized to a fixed size of 1280x1280 pixels. YOLOv5s employs anchor boxes and predicts bounding box coordinates, objectness scores, and class probabilities in a single forward propagation through the network. The performance of the YOLOv5s model is evaluated using metrics such as precision, recall and mAP@0.5. The detected regions of interest were cropped using the predicted bounding box coordinates and saved for text recognition.

Fig. 2. Detection result using YOLOv5

2.2 Nameplate Text Recognition

This section introduces recognizing text information from the detected nameplate area. When an image is input, the learned yolov5 model crops the image as bounding box coordinates after detecting nameplate. Tesseract OCR was adopted to recognize the text of the separated image after image preprocessing.

2.2.1 Preprocessing. For each cropped nameplate image, the following preprocessing steps were performed:

a. Color conversion: The input image is first converted from the RGB color space to grayscale for simplifying the image by retaining only the intensity information and discarding color information.

b. Inversion: The grayscale image is then inverted. This operation flips the pixel values, effectively turning dark regions into light regions and vice versa. Inverting the image can improve the performance of the following thresholding step, especially for images with light text on a dark background.

c. Thresholding: The inverted grayscale image applies binary thresholding The Otsu's thresholding method is used to automatically determine the optimal threshold value. This process helps to separate the text regions from the background, which can enhance the accuracy of the OCR engine.

2.2.2 Text Recognition. Tesseract OCR is applied to the preprocessed nameplate images to extract the text information. The performance of Tesseract OCR on the nameplate images is evaluated by compared to ground truth.

3. Results

3.1 Nameplate Detection

The results show that the YOLOv5s-based nameplate detection model can effectively detect different type of nameplate from natural scene with precision of 96.8%, recall of 97.1% and mAP@0.5 of 97.9%. The model showed very high performance compared to learning with less data than is required for deep learning., and it is also successful to simultaneously detect multi-objects from images taken in various real-world environments.

Table 1: Performance Metrics of the YOLOv5s-based Nameplate Detection

Performance Metric	Value (%)
precision	96.8
recall	97.1
mAP@0.5	97.9

3.2 Nameplate Text Recognition

The result of text recognition is precision of 79.0%, recall of 84.7% and F1 score of 81.7%. However, this result is for 119 detected objects which Tesseract OCR is applied. For several nameplates, the text was not recognized using Tesseract OCR. In the case of component nameplates, when the nameplates are cropped after detection, the image size becomes too small, making it difficult for Tesseract OCR to operate properly. In addition, the case of false positives was very high, most of which recognized bolts for fixing nameplates as text such as ‘@’, ‘©’, and ‘(4’. Table 3 shows an example of false positive detection in the test set. To improve this, it is necessary to apply a process of detecting a text area of a nameplate before OCR in the future.

Table 2: Performance Metrics of the Text Recognition

Performance Metric	Value (%)
precision	79.0
recall	84.7
F1 score	81.7

Table 3: Example of False Positive Detection

Result of Each Stage of Text Recognition	
Cropped Nameplate Image	
After Preprocessing	
Ground Truth	ANNUNCIATOR 3-598-J-UL01
OCR Result	, ANNUNCIATOR © 3-598-J-UL01 ©

4. Conclusions

In this paper, a deep learning approach is proposed for detecting nameplates and recognizing text using YOLOv5 and Tesseract OCR. Experimental results demonstrated that the proposed approach achieves high performance. The proposed, which recognizes nameplate information through images, has a potential of entering various applications and businesses in the future.

REFERENCES

[1] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in neural information processing systems, 25, 1097-1105.
[2] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
[3] Smith, R. (2007). An overview of the Tesseract OCR engine. In Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Vol. 2, pp. 629-633.
[4] Tzutalin. LabelImg. https://github.com/tzutalin/labelImg
[5] Simard, P.Y., Steinkraus, D., and Platt, J.C. Best practices for convolutional neural networks applied to visual document analysis. In Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings., pages 958-963. IEEE, 2003.
[6] Bochkovskiy, A., Wang, C.Y., and Liao, H.Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934, 2020.
[7] Wang, C., Xu, J., and Chunjing, X. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. arXiv preprint arXiv:1911.11929, 2019.
[8] Liu, S., Qi, L., Qin, H., Shi, J., and Jia, J. Path Aggregation Network for Instance Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 8759-8768, 2018.