Effective Detection of Materials in Construction Drawings using YOLOv4-based Small Object Detection Techniques

Jiwoo Sim, Heejo Woo, YunHwan Kim*, Eung-Tae Kim
Dept. of Electronic Engineering. Korea Polytechnic University, Siheung, KOREA
WooSung Steel. Inc, Seoul, KOREA*
2020811010@kpu.ac.kr, wq8336@gmail.com, ceo@woosungsteel.com*, etkim@kpu.ac.kr

***Abstract*— Since the quantity surveying of the materials marked on construction drawings is conducted manually, it is very time-consuming and causes problems such as incorrect calculation transactions. So, a fast and accurate AI-based automatic quantity surveying system is required. In order to accurately detect steel materials in construction drawings, we propose data augmentation techniques and spatial attention modules for improving small object detection performance based on YOLOv4. Experimental results show that the proposed method increases the accuracy and precision by 1.8% and 16% respectively compared with the conventional YOLOv4.**

Keywords—Construction Drawings, Text Recognition, YOLO, Data Augmentation, Spatial Attention

I. INTRODUCTION

Artificial intelligence-based systems have been developed in various fields and have been applied to various areas. Although artificial intelligence-related technologies are being applied in the construction industry, they are still in the early market phase. In particular, the system of trading steel materials marked on the construction drawings is currently being conducted manually. So, the work is very time-consuming and various problems arise such as the difficulty of accurate quantity surveying and false calculation transactions. These problems can be solved through the development of an artificial intelligence-based automatic quantity surveying system as shown in Fig.1, that involves an auction system that automatically identifies the type and quantity of steel materials to swiftly and accurately determine the construction unit costs such as the steel material costs provided by steel suppliers. However, studies on artificial intelligence-based construction drawings are still limited to recognizing symbols in drawings [1]. Therefore, in order to develop an automatic quantity surveying system based on artificial intelligence, it is necessary to study the detection and recognition of steel materials in construction drawings.

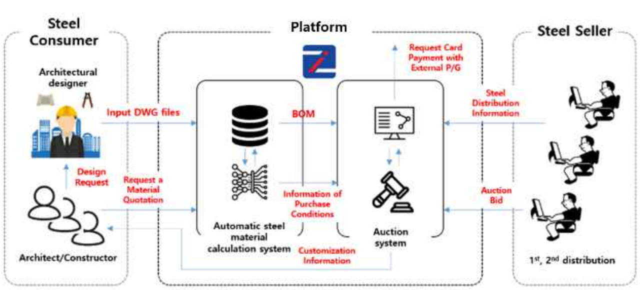


Fig 1. Auction System based on Automatic Quantity Surveying of Steel Materials

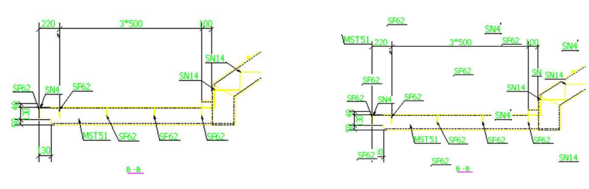
To implement the detection part of steel materials in the automatic quantity surveying system, detection should be speedy and accurate for small objects like texts of steel materials in large construction drawings.

II. PROPOSED METHOD

In order to learn large-sized construction drawing images rapidly, we use a YOLOv4 model [2] which enables learning with a single GPU. While 1-stage detectors such as YOLO have a quick processing speed, they have the disadvantage of poor in detecting small objects. The detection performance improvement is crucial in detecting small objects, which are mostly texts of steel materials in the construction drawings. Therefore, we propose data augmentation technique to improve the learning effect and YOLOv4-based detection technique which apply the spatial attention module to detect small objects as shown in Fig. 3.

First, we used the data augmentation technique [3], a type of replication method. The original method involves the replication of objects of size 32x32 or less without overlapping in random locations based on the COCO dataset. In order to apply this technique to a construction drawing, we must consider cases where texts of steel materials are shown to overlap by replicating every object since most texts of steel material objects are of size 32x32 or less. We can then secure diversity of location and case by allowing partial overlap of objects as shown in Fig 2(b) under the condition that most objects are not overlapping with one another.

Second, the spatial attention module [4] is applied before transferring the extracted features to the detection stage, YOLO as in Fig 3. The original spatial attention module ensures large receptivity by applying 7x7 convolution to channels compressed by average pooling and max pooling as shown in the diagram depicting the spatial attention module in the bottom left of Fig 3. In Equation (1), s means spatial attention, M_s refers to the calculated value of spatial attention module, σ refers to the sigmoid function, F^s_{avg} refers to the average pooling value, and F^s_{max} refers to the max pooling value.



(a) (b)

Fig 2. (a) the enlarged image of a section of the original construction drawing, (b) the construction drawing obtained from applying the data augmentation technique in the same area as (a)

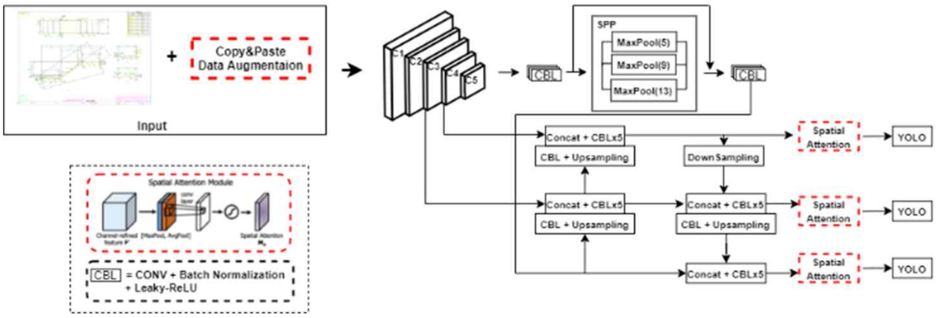


Fig 3. Full Block Diagram of the Proposed Learning Model Structure

$$M_s(F) = \sigma\left(f^{3x3}([AvgPool\ (F); MaxPool\ (F)])\right) \qquad (1)$$
$$= \sigma\left(f^{3x3}\left(\left[F^s_{avg};\ F^s_{max}\right]\right)\right)$$

III. Experimental Results

The construction drawing used in the experiment has a size of 1584x1224, and there is a lot of information besides the name of the steel material to be detected as shown in Fig 2(a). 719 sheets were used in learning, and 80 were used in testing. We tested the simultaneous application of the data augmentation technique and the spatial attention module that were proposed to improve the ability to small-sized text detection of steel materials.

Table 1 shows the comparative results of applying the YOLOv4 model, the spatial attention module, the data augmentation technique, and when the latter two methods were both applied. We found that the precision and the accuracy drastically increased when only the spatial attention module was applied and when the convolution calculation size after average pooling and max pooling was modified to 3x3. When the data augmentation technique of the replication method was applied, the precision increased and the accuracy also increased by 1%. Therefore, when both modules were simultaneously applied, the accuracy increased by up to 1.8%, and the precision was maintained at more than 90% with the highest precision value at 94.8%.

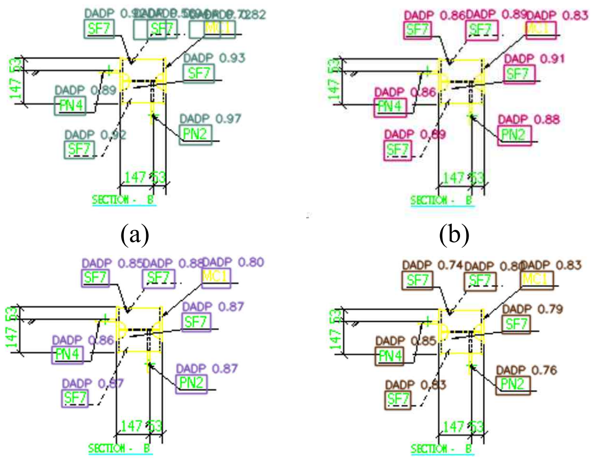
Fig 4 is an enlargement of some of the detection results of the actual drawing image. Fig 4(a) shows the results of applying only YOLOv4, which resulted in duplicated detection. The results of applying the data augmentation method and the spatial attention module are displayed in Fig 4(b) and Fig 4(c), respectively. Fig 4(b) shows that while all objects were detected, the point of detection has slightly deviated. The detection results in Fig 4(c) were stable, but the individual accuracy has slightly decreased. Fig 4(d) shows the results of simultaneous application of both the data augmentation method and the spatial attention module. We found that this method produced the most stable object detection results as there were no objects undetected or detected in duplicate, and the individual accuracy for each object has also increased.

IV. Conclusions

Effective detection method based on Yolov4 is proposed to accurately detect small size of texts of steel materials in large construction drawings. The proposed method modified and applied the data augmentation technique of the replication method and the spatial attention module. The experimental results show that the proposed method has better performance than the conventional Yolov4 method.

TABLE I. Results of Applying the Spatial Attention Module, the Data Augmentation Method, and the Both

Methods	*Precision*	*Recall*	*mAP@.5:.95*
YOLOv4	***77.9***	***99.8***	***67.0***
SA-7x7	***83.4***	***99.8***	***66.1***
SA-3x3	***94.8***	***1***	***67.1***
Aug x2	***82.2***	***1***	***68***
Aug x3	***84.2***	***1***	***68***
SA+Aug x2	***94.5***	***1***	***68***
SA+Aug x3	***94.8***	***1***	***68.8***



(a) (b) (c) (d)

Fig 4. (a) Detection results using YOLOv4, (b) Results of applying the data augmentation technique with replication twice, (c) Detection results of applying the spatial attention module (d) Detection results of applying both of (b) and (c)

References

[1] Eunseop Yu. "Approach of object recognition from image format engineering drawings using deep learning." Master's thesis, KyoungBuk University, 2019.

[2] Alexey Bochkovskiy, Chien-Yao Wang, and HongYuan Mark Liao. YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, 2020.

[3] M. Kisantal, Z. Wojna, J. Murawski, J. Naruniec, and K. Cho, "Augmentation for small object detection," arXiv preprint arXiv:1902.07296, 2019.

[4] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. CBAM: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), pages 3–19, 2018.