Transactions of the Korean Nuclear Society Autumn Meeting
Changwon, Korea, October 20-21, 2022

Information Management for Nuclear Decommissioning: Synthesizing Text with Drawings

Hokyun Kim[a], Minhyuk Jeong[b], Sihyun An[b], Juyub Kim[c],
Hyungi Byun[c], Doo-Yong Lee[c], Jinu Kim[b], Kwangwon Ahn[b, *]

[a]*Graduate School of Future Strategy, KAIST, Daejeon 34141, South Korea*
[b]*Department of Industrial Engineering, Yonsei University, Seoul 03722, South Korea*
[c]*FNC Technology Co., Ltd., Institute of Future Energy Technology, Yongin-si, Gyeonggi-do 16954, South Korea*
*Corresponding author: *k.ahn@yonsei.ac.kr*

1. Introduction

The nuclear power plant (NPP), one of the most delicate and advanced man-made structures, has a complex structure and no margin for error, as any radioactive leak could affect public health. As a result, any NPP's decommissioning procedure should adhere to the "as low as reasonably achievable" principle. The multi-agency radiation survey and site investigation manual (MARSSIM) requires a thorough investigation of big data related to the respective NPP for decommissioning. This is an essential step, known as site characterization, in any NPP decommissioning project.

In case of potential contamination, humanity has previously relied entirely on the labor of professional personnel to locate and mark the location of the incident. As a result, the resource consumption and opportunity costs for historical site assessment (HSA)[1] have been substantial. We propose a deep neural network (DNN)-based framework to both alleviate the burden of human labor and accelerate the overall execution speed of the detection and marking process. On NPP drawings, our goal is to identify and mark all boundaries of areas contaminated by the corresponding event. Specifically, our framework is conceptually based on [1], in which images are synthesized according to text; in our case, it refers to the boundaries within an NPP drawing that corresponds to a specific radioactive incident. Our model structure is based on [2], which introduces a methodological innovation aimed at lowering computational costs when training and applying DNN-based frameworks, particularly image-based framework.

2. Method and Results

The decommissioning of an NPP begins with HSA. The HSA procedure utilizes i) reviews on design documents, ii) operational history, iii) official reports on nuclear contamination events, and iv) NPP drawings. As previously stated, NPPs have a complex structure with little room for error. This is comparable to traditional image-related DNN-based projects with the goal of generalization. This is demonstrated in [2], where the goal of benchmark datasets is to create general and widely applicable ensembles of image data. In such projects, state of the art results typically fall between 80% and 95% margin. In contrast, NPP experts have advised that our framework should be scored even higher.

2.1 Framework

We aim at synthesizing the image representation of a room in an NPP drawing with semantic representation, augmented within text. Our semantic representation will be hereafter referred to as a query. Each query is hypothesized to contain information on i) what structures, systems, and components were contaminated and ii) where the corresponding figure is in the drawing. Our model only uses text data to match the query with the corresponding image representation. Since queries, which are semantic representations and thus solely text data, are directly matched with and converted to their corresponding image representations, we refer to our model as an object conversion system (OCS). OCS is an end-to-end framework that eliminates the need for postprocessing methods on both input and output, as well as additional framework configurations.

2.2 Evaluation Criteria

We assess the accuracy of our text model using the intersection over union (IoU) metric calculated using the Jaccard index [3]. Because the output consists of four coordinates, we have tentatively concluded that IoU is the appropriate metric for evaluation. IoU is calculated by dividing the intersection of the bounding boxes for ground truth data and predicted data by their union area, which can be expressed as:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}, \tag{1}$$

where A denotes the area of ground truth data and B denotes the corresponding predicted area.

2.3 Data Selection and Preprocessing

In contrast to most DNN-based projects, where related datasets for the task are mostly open source,

---

[1] HSA is a preliminary site investigation process where potentially contaminated areas are to be visually distinguished according to every available data on the according NPP.

related data for NPPs are frequently classified and unavailable for public use. Accordingly, in order to train and test our DNN-based module, we improvise by selecting and utilizing data that is similar to that of actual NPP drawings. NPP expert personnel manually collected our dataset via web image search.

NPP drawings are classified mainly in two categories: room number (RN) drawings where each room and its room number are marked and a corresponding general assessment drawing, which contains information on the names of structures, systems, and components within each room. We use open materials analogous to RN drawings to simulate text-to-image synthesis using OCS. Fig. 1 depicts a sample from our dataset.

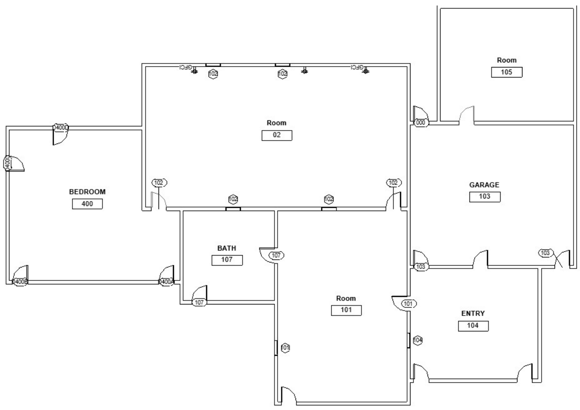

Fig. 1. An example of an RN drawing. The drawing, in which we will map the outputs, depicts each compartment that could be potentially contaminated.

The RN labels are supposed to be raw pixel coordinates per image, which are then used as ground truth labels for each room in the RN drawing, corresponding to A in the equation (1). We enter our query corresponding to A into OCS, and it returns B. Finally, we evaluate OCS results based on equation (1).

2.4 Object Conversion System

We use solely text data to infer the image representation indicating potentially contaminated areas. To extract relevant information from queries, we first tokenize and vectorize semantic representations into numerically processable representations, a process known as word embedding [4].

OCS is made up of an embedding layer, convolutional layers with residual connections [2], and four head layers that infer the coordinate value of all four vertices. Most word embedding frameworks use the recurrent layer-based DNN model (RNN) [4]. However, to reduce the computational load of OCS, we use a convolutional layer-based DNN model (CNN) instead of RNN. The computation load of RNN is exponentially proportional to the size of the vector sequence, whereas CNN is only linearly proportional to. As a result, we concluded that implementing CNN will scale efficiently to longer text lengths and query sizes. We also use an innovative method from [2] in which the residual connection is used instead of the pure sequential connection. The implemented innovation is as follows:

$$H(x) = F(x) + x, \tag{2}$$

where $H(x)$ denotes the output of residual connections, $F(x)$ denotes the weighted vector sequences, i.e., layer block results, and x denotes the entering vector sequences before passing through each layer block.

Following convolutional layers, we facilitate the layer connections using a global average pooling layer and stack fully connected dense layers for calculating outputs. The coordinates of four vertices (i.e., minimum x- and y-coordinate, maximum x- and y-coordinate) are estimated by the head layer. In contrast to other object detection tasks, such as [5], we must map a single text sequence to a single contaminated region. Therefore, following fully connected dense layers, four identical dense layers are connected in parallel.

2.5 Results

Fig. 2 depicts the detected region for one of our samples using our model. OCS returns an IoU value greater than 99.8%, which is sufficient for HSA.

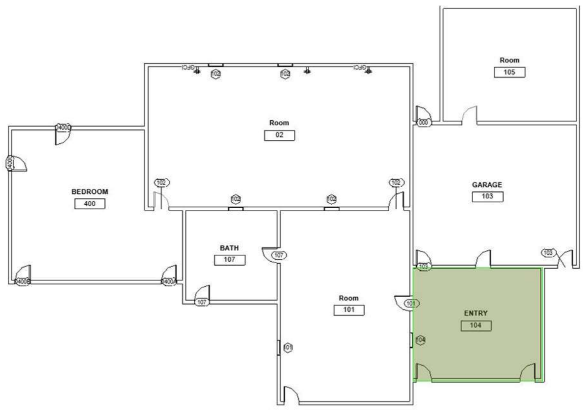

Fig. 2. RN drawings annotated with OCS-detected regions.

We summarized the results of our trained model using 7 different input text queries in Table 1. It has been demonstrated that 6 out of 7 IoU results achieved over 99%. This means that our model effectively captures and converts the latent information in each query to image representation using only text input. However, we only achieved approximately 90% accuracy for Query #7 (Room name: Bedroom 400). This is due to the limitations of our labeling procedure; our model can only label rectangular images, whereas the corresponding image representation of Query #7 does not fall into this category. This must be addressed in future studies if the model is to be used properly in NPP decommissioning projects.

Table 1. Test results of the trained model

Room name	Query	IoU (%)
Room 105	룸 105 번에서 오염 발생	99.30
Garage 103	개러지 103 번에서 오염	99.71
Entry 104	엔트리 104 번에서 오염 발생	99.84
Room 02	룸 02 번에서 오염 발생	99.91
Room 101	룸 101 번에서 오염 발생	99.92
Bath 107	배쓰룸 107 번 에서 오염 발생	99.92
Bedroom 400	베드룸 400 번에서 오염 발생	90.20

3. Conclusion

We propose a method for automating and lowering the cost of an extensive site characterization process for decommissioning NPPs. We propose OCS, a model capable of converting only text input to image outputs. Our research would serve as a backbone for implementing lightweight, yet sufficiently accurate procedures for decommissioning NPPs. Not only our research applies to HSA, but it also applies to the entire MARSSIM site characterization procedure for decommissioning NPPs.

In the future, we should focus on less accurate results, such as Query #7 from Table 1. According to experts, NPPs typically revolve around circular structures, emphasizing the importance of a labeling method capable of annotating non-rectangular images. We leave it to future research to implement improvisations capable of labeling various forms of image representation systematically.

ACKNOWLEDGEMENT

This work was supported by the Technology Innovation Program ATC+ (20014125, Development of Intelligent Management Solution for Nuclear Decommissioning Site Characterization) funded by the Ministry of Trade, Industry & Energy (MOTIE, Republic of Korea)

REFERENCES

[1] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, H. Lee, Generative adversarial text to image synthesis, In *Proceedings of the 33rd International Conference on Machine Learning*, pp. 1060–1069, 2016.
[2] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 770–778, 2016.
[3] M. A. Rahman, Y. Wang, Optimizing intersection-over-union in deep neural networks for image segmentation, In *International Symposium on Visual Computing*, pp. 234–244, 2016.
[4] S. Lai, K. Liu, S. He, J. Zhao, How to generate a good word embedding, *IEEE Intelligent Systems*, 31(6), pp. 5–14, 2016.
[5] J. Redmon, S. Divvala, R. Girshick, A. Farhad, You only look once: Unified, real-time object detection, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788, 2016.