ICIC Express Letters
Volume **16**, Number **6**, June **2022**

ICIC International ©2022 ISSN 1881-803X
pp. 621–628

A STUDY ON THE MICROORGANISMS IMAGE RECOGNITION ASSIST SYSTEM BASED ON DEEP LEARNING

Hiroshi Shiratsuchi[1], Misato Furuta[2], Yuji Wakamori[1], Kozo Horiuchi[3] and Takanori Matsuzaki[1]

[1]Graduate School of Humanity-Oriented Science and Engineering
Kindai University
11-6 Kayanomori, Iizuka-shi, Fukuoka 820-8555, Japan
{ sira; takanori }@fuk.kindai.ac.jp; wakamori@iis.elec.fuk.kindai.ac.jp

[2]Cowsy Co., Ltd.
488-1 Kayanomori, Iizuka-shi, Fukuoka 820-0011, Japan
envi@cowsy.jp

[3]Gururi Co., Ltd.
2-1-6 Uomachi, Kokurakita-ku, Fukuoka 802-0006, Japan
kozo.horiuchi@gururi.co.jp

Received October 2021; accepted December 2021

Abstract. *This paper aims at a microorganisms image detection system that uses machine learning to estimate the status of aerobic microorganisms in activated sludge, which is necessary for the stable management and operation of a purification facility for factory wastewater. In other words, if the proposed system supports the detection of the type and number of microorganisms in the aeration tank, the condition of the purification facility is easy to estimate. The GoogLeNet of deep learning models trained by datasets with Augmentation such as a padding process is proposed. From the simulation results, the recognition rate for the validation data was a good result in 81.36%.*
Keywords: Microorganisms, Wastewater treatment facility, Machine learning, Convolutional neural networks

1. **Introduction.** There are regulation values for the quality of water discharged from factories defined by the Ministry of the Environment in Japan [1]. For this reason, factories purify the water so that it does not exceed the effluent standards before releasing it to the river. Such wastewater treatment facilities use aerobic microorganisms called activated sludge to decompose and treat the pollutants in the wastewater [2]. A high level of expertise and experience is required for the stable operation of facilities. However, the problem is that the time and cost necessary to inherit these are limited.

For the operation and maintenance of the current wastewater treatment facilities, technical employees visit the installation site of the treatment facilities once a week to once a month to conduct the below investigations and adjustments. First, each facility is fine-tuned based on visual confirmation and data obtained from water quality and various sensor devices (pH, dissolved oxygen, oxidation-reduction potential, activated sludge suspended solids and others). The validity of these operations on-site is confirmed by analyzing microscopic observation of the microorganisms from bringing back the activated sludge collected by facilities. If the on-site response is inadequate, simple manipulation, such as valve adjustment, is requested to the local personnel who do not have expertise in the facility as a remote response. At the next visitation of technical employees, appropriate adjustments operate. Therefore, if the status of microorganisms is analyzed on-site, a quick and accurate response by technicians is possible.

DOI: 10.24507/icicel.16.06.621

As mentioned above, the types and numbers of microorganisms appearing in activated sludge are correlated with the condition of the wastewater treatment facility. However, the estimation accuracy is not stable due to differences in the microorganisms identification ability of technical employees in actuality. For this reason, several methods have been proposed to automatically detect target microorganisms, such as analysis methods using phase matching of microscope images [3] and classification of the processed images by SVM (Support Vector Machine) [4]. However, the microorganisms detection by these methods has not been sufficiently accurate.

In machine learning, the deep learning model trains using input and teacher pairs called datasets. In other words, select the best deep learning model to extract the trends and features of the target problem, and this model trains by the training data of the datasets. The model obtained in the above process can recognize, identify, classify, and predict.

This paper aims to develop a microorganisms image detection system that supports the inheritance of expertise in the operation and maintenance of wastewater treatment facilities. Specifically, sludge material from a facility called an aeration tank, where aerobic microorganisms decompose organic substances in wastewater, is collected to collect microorganisms images by the microscopic. In a process called annotation, extracted region of the target microorganisms from the microbe image assigns as a label such as a microorganism name. Training CNNs (Convolutional Neural Networks) usually requires large datasets of thousands to tens of thousands [5]. In addition, microbial images often contain inverted or rotated images due to their size and shape. Therefore, we will also consider a padding method called Augmentation [6] for datasets by processing such as image rotation, scaling, inversion, and translation.

As described above, the proposed system will expect to use effectively as technical employees' training and support software. In the following, the outline of the wastewater treatment facility is described in Section 2. An Augmentation method is introduced in Section 3 that the images of target microorganisms cropped by annotation were converted as an extension of the affine transform. Section 4 introduces GoogLeNet, which is one of the representative examples of deep learning models. Then the results of actual training are also described. Section 5 summarizes the paper.

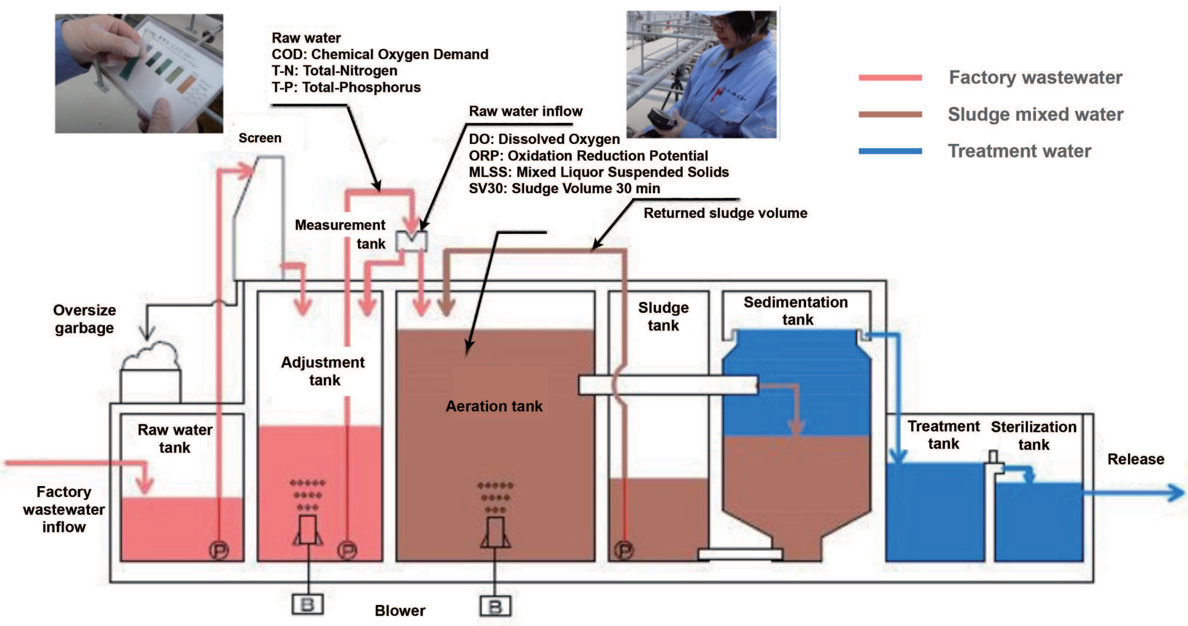


FIGURE 1. (color online) Overview of wastewater treatment facilities

2. **Overview of Wastewater Treatment Facilities.** Factory wastewater varies greatly depending on the type of operation. In this section, we will use a food manufacturing plant as an example. The wastewater in these factories does not contain toxic substances, but it often contains substances that may cause environmental pollution [7]. Figure 1 shows the diagram of the wastewater treatment facility based on the standard activated sludge method treating wastewater from a food manufacturing plant. In the figure, the aeration tank uses an aerobic microbial treatment system called activated sludge. In addition, due to the nature of food production, the amount and quality of water in the factory wastewater varies from day to day because of the raw materials handling, the processed items manufactured, and the production volume changes dynamically with the seasons. Therefore, the operation and maintenance of the wastewater treatment system have a property of the frequent need for fine adjustments [2].

Extracted from the design document as the rated capacity of the wastewater treatment facility is as the following items: (a) Design water volume, (b) Design water quality (BOD[1]), (c) Treating method (d) Capacity of biological treatment tanks of Aeration, Pre-aeration and BF, (e) Sedimentation tank volume, (f) Water area of sedimentation tank, (g) Membrane treatment capacity[2], (h) BF thread length[3].

The operation of wastewater treatment facilities in both aeration and sedimentation tanks is critical to pollutant treatment performance. Therefore, the facility management indicators as Table 1, essentially define as operation management items focusing on two tanks. Therefore, the operation management items calculated based on the numerical values measured in the field survey of the facility are evaluated and addressed.

Table 1. List of the operation management items

	Operation/Maintenance	Indicators	Description
1	Influx of water	Rated capacity	Indicate water volume load.
2 3	Nitrogen to BOD Phosphorus to BOD	Recommended Recommended	Indicate the nutrient balance in wastewater. An example of an ideal ones is BOD:T-N:T-P = 100:5:1.
4 5	BOD volume load Load of BOD-SS	Rated capacity Recommended	Amount of BOD in relation to the amount of volume and microorganism in the aeration tank per day.
6	SVI	Recommended	Sedimentation property of activated sludge.
7	Diffused air status	Recommended	Remaining oxygen in the aeration tank after chemical oxidation of pollutants and respiration of microorganisms.
8	Sedimentation tank residence time	Rated capacity	Time for solid-liquid separation.

The status of decomposition and treatment of pollutants in factory wastewater depends on the presence of microorganisms in the treatment equipment. There is a relationship between the condition of the processing equipment and the type and quantity of microorganisms present, which is well known. Therefore, these indicators are significant factors

[1]Abbreviation for Biochemical Oxygen Demand, an indicator of water pollution. BOD is the amount of oxygen required for aerobic microorganisms to decompose organic substances in water.

[2]The treatment capacity of the membrane filtration system for solid-liquid separation used in the membrane bioreactor method. Compared to the standard activated sludge method, the membrane filtration system can be installed in a smaller area and can significantly remove turbidity from the treated water.

[3]The length of the microorganism retention material that adsorbs BF (Biofringe) activated sludge, in other words, is referred to as BOD treatment capacity.

in determining the operation and maintenance operations of the wastewater treatment facility. However, it is difficult to observe them at the site due to the equipment, and the activated sludge is usually taken back and observed under an electron microscope. The time required to determine the water quality of the aeration tank may range from 15-30 minutes for skilled technicians to more than 2 hours for an inexperienced technician. Furthermore, even among them, differences in identification ability can lead to differences in the inferred results of the state of the processing equipment.

3. **Identification of Microorganisms Appearing in Wastewater Treatment Systems.** This section covers microorganisms that occur in aeration tanks at wastewater treatment facilities. Specifically, the following process is developed from the gathering of microorganism image data to identification by a deep learning model.

(i) Gathering of microscopic image data of aeration fluid collected from the wastewater treatment system.
(ii) Select microorganisms for identification.
(iii) Annotation of the target microorganisms within the microscopical images.
(iv) Dataset generates for deep learning and padding by Augmentation.
(v) Training of the deep learning model and validation of recognition capability.

The following section describes these details.

3.1. **Gathering of microscopic image data from activated sludge.** The activated sludge was collected from 38 wastewater treatment facilities in food manufacturing plants. A total of 5,736 images were extracted from collected samples by the microscope. These image data contain a wide variety of microorganisms. Twenty microorganisms shown in Table 2 were selected for the necessity to estimate water quality.

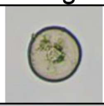
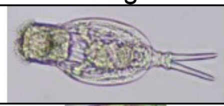
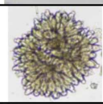
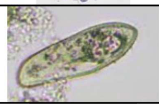
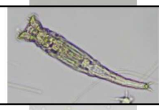
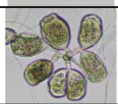
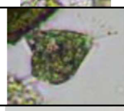
In the table, microorganisms such as Arcella and Centropyxis appear in large numbers when nitrification in the tank progresses and the pH of the treated water decreases. In addition, Lecane and Lepadella have the characteristic of appearing when the influent water concentration is low and the load is extremely low. Moreover, Epistylis and Euglypha appear in large quantities when the facility is under low load and sludge demolition is in progress. Other microorganisms also vary in the types and populations that appear depending on the load of the aeration tank or wastewater treatment facility.

3.2. **Annotation and Augmentation of biological images.** Wastewater conditions greatly rely on the presence or absence of these microorganisms, their numbers, and their combination. Therefore, if determining the number of each microorganism per unit area, the condition of the wastewater treatment facilities will be possible to estimate. Then, the acquired image data is annotated based on polygonal approximation and labeling of the area containing the microorganisms to detect using an application called Microsoft VoTT (Visual Object Tagging Tool) [8]. The above steps generate the fundamental dataset required for deep learning.

However, in general, tens of thousands of image data are required to achieve sufficient recognition accuracy in deep learning. Therefore, the fundamental datasets are processed images affine transformation such as rotating, scaling, flipping, and parallel translation. In other words, image rotations assign as a random number in the range of $\pm 5°$ in $15°$ increments from $0°$ to $345°$. Furthermore, the image is scaled to maintain the aspect ratio of the image data, flipped a probability of 50%, and parallel translated with a random number less than 5% of the image size. This procedure is shown as Figure 2. As a result, the number of fundamental datasets increases 24 times.

From the above procedure, 41,064 datasets were prepared for training data by deep learning, 664 validation data as a measure of learning progress, and 231 test data for performance evaluation of the completed deep learning model. Here, the validation data is a part of the dataset that is not directly used for training but is regularly presented to

Table 2. List of target microorganisms

No.	Name	Image	No.	Name	Image
1	Arcella		11	Lecane	
2	Aspidisca		12	Lepadella	
3	Blepharisma		13	Opercularia	
4	Carchesium		14	Paramecium	
5	Centropyxis		15	Peranema	
6	Chaetonotus		16	Philodinidae	
7	Chaetospira		17	Prorodon	
8	Epistylis		18	Pyxidicula	
9	Euglypha		19	Tokophrya	
10	Euplotes		20	Vorticella	

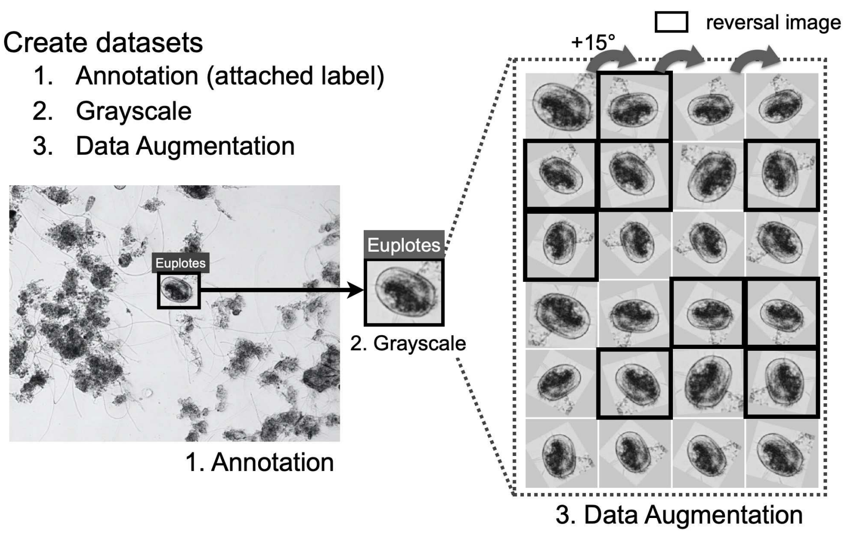


Figure 2. Overview of dataset generation

the model as a criterion to measure the learning progress. Test data is completely unused data not used for training or validation. Note that the verification data and test data are not included of images transformation by Augmentation.

4. **Learning by Convolutional Neural Networks (CNNs).** The Convolutional Neural Networks (CNNs) are widely used for regions such as image recognition. The CNNs realize image feature extraction by repeatedly applying a convolutional layer and a coupling layer to input data such as a two-dimensional image. Then, image identification enables by learning in the fully connected layer. One of the typical implementations of this convolutional neural network is a deep learning model called GoogLeNet proposed by Szegedy et al. [9]. This model won the ILSVRC (ImageNet Large Scale Visual Recognition Challenge), a contest held in 2014 that competes for identification accuracy and speed on ImageNet, one of the image recognition datasets. The GoogLeNet has a multiple Inception Module consisting of convolutional layers of different sizes, such as 1×1, 3×3, and 5×5, and a pooling layer of 3×3. This inception module reduces the number of feature maps by decreasing the dimensionality of the input data, resulting in more efficient training compared to conventional CNNs. In this section's GoogLeNet, there are 22 layers included of nine Inception Modules combination.

4.1. **Learning results by GoogLeNet.** The GoogLeNet of deep learing models in the above paragraph was performed on nVidia DIGITS[4] using the training datasets consisting of the image size 256×256 pixels, including 41,064 images as training data and 664 images as validation data. As a result of training, the recognition rate for the validation data was a good result in 81.36% (Figure 3). Furthermore, Table 3 shows the results of measuring the recognition rate for each category using the 231 test data. The vertical column is the category number of the microorganism in Table 2 to which the test data set belongs, and each column represents the number of numbers identified by GoogLeNet. This table is called a confusion matrix, and the diagonal columns where the vertical and horizontal category numbers match indicate that the identification is correct. The rightmost column shows the percentage of correct answers for each category. The correct response rate

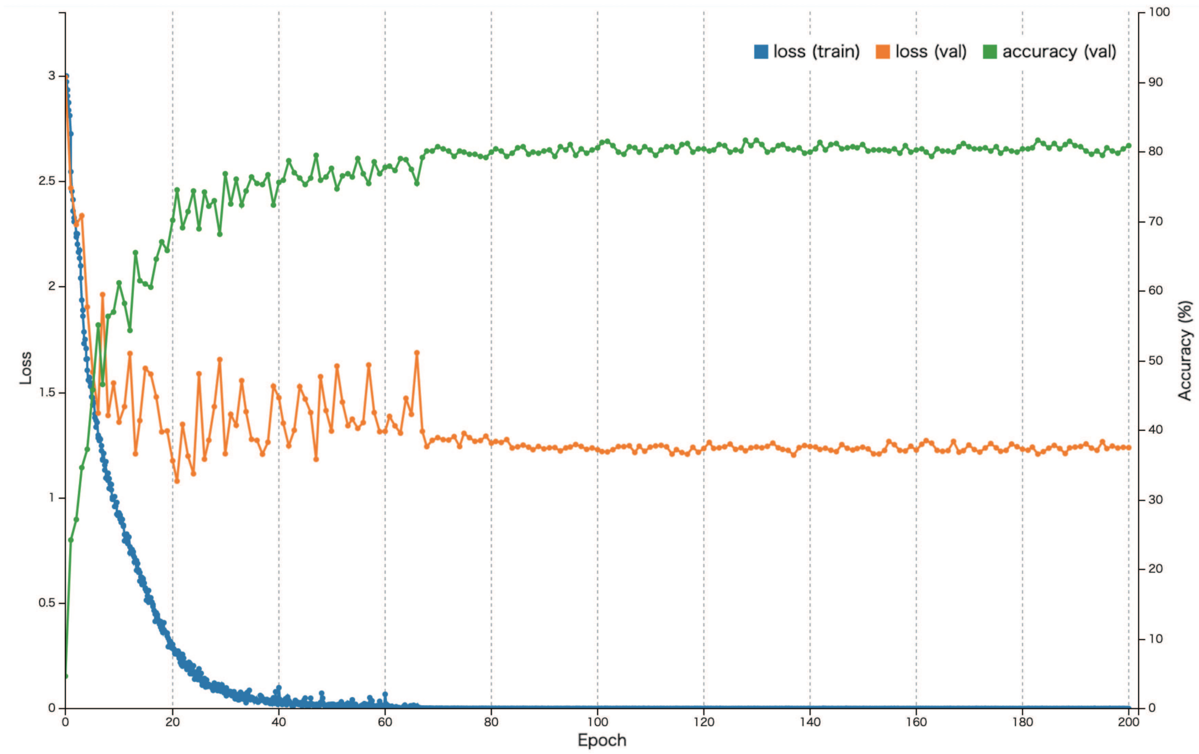


FIGURE 3. (color online) Transitions of loss and accuracy for training and validation data

[4]https://developer.nvidia.com/digits

TABLE 3. Confusion matrix of learning results

Predicted Class

Actual Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	accuracy
1	8	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	88.9%
2	0	39	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	97.5%
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
4	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	25.0%
5	0	0	0	0	18	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	94.7%
6	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	1	0	0	0	1	50.0%
7	0	0	0	0	0	0	14	0	0	0	0	1	0	0	0	0	0	1	0	0	87.5%
8	0	0	0	0	0	0	0	8	1	0	0	0	1	1	0	0	0	0	0	4	53.3%
9	1	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	1	0	0	83.3%
10	0	0	0	0	0	0	1	0	0	3	0	0	0	0	0	0	0	0	0	0	75.0%
11	0	0	0	0	0	0	0	0	0	0	18	2	0	1	0	1	0	1	0	0	78.3%
12	0	0	0	0	0	0	0	0	0	0	1	14	0	1	0	0	0	0	0	0	87.5%
13	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	100%
14	0	0	0	0	0	1	0	0	0	0	0	0	0	5	0	1	0	0	0	0	71.4%
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	100%
16	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	16	0	0	0	0	84.2%
17	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	1	0	0	84.6%
18	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	6	0	0	54.6%
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	100%
20	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	19	95.0%

was a good result in 80% to 100% for most categories. From the table, we can see that the identification rate of 4 (Carchesium) is insufficient. For example, the total number of images in categories 3 (Blepharisma), 6 (Chaetonotus), 10 (Euplotes), 13 (Opercularia), and 19 (Tokophrya) is also little, and the reliability of the recognition rate is considered low. However, the overall recognition rate of 83.73% for test data is almost the same as 81.36% for the validation data. From these results can conclude that the proposed microbial detection system has sufficient discrimination capability.

4.2. **Learning results by other models.** In this section, the results of training with ResNet [10] and InceptionResNetV2 [11] as model structures are described. The simulation was implemented as transfer learning using Keras [12] instead of nVidia DIGITS. In addition to the Augmentation method described in Section 3.2, we also compared and examined the generator function using *keras.ImageDataGenerator()*[5] on the Keras APIs. The results of these simulation experiments are shown in Table 4. However, any recognition rate for the model structures and Augmentation methods could not be surpassed by the GoogLeNet ones.

TABLE 4. Simulation results using several CNN models

Model name	Augmentation	Recognition rate
ResNet50	Keras API	68.5%
ResNet50	proposed	69.2%
ResNet152	Keras API	76.5%
ResNet152	proposed	75.3%
InceptionResNetV2	Keras API	77.5%
InceptionResNetV2	proposed	80.2%

[5]https://keras.io/api/preprocessing/image/

5. **Conclusion.** This paper proposes a microorganisms image detection system that supports the inheritance of expertise in the operation and maintenance of wastewater treatment facilities. The proposed detection system was able to identify 20 microorganisms necessary for the determination of water quality with high accuracy. By using this system, microorganisms identification is possible regardless of the skill level of the technician. However, our proposed detection system may not function properly depending on the image containing a lot of information other than microorganisms or the scale magnification of the microscope image. Since a low-magnification microscope image includes a wide range, the proposed system fails to identify when multiple or non-target microorganisms appear or dust is mixed in a large amount.

To solve these problems, the YOLOv3 [13, 14], another deep learning model, considers using instead of the GoogLeNet. However, in this case, the training process requires more image datasets. For the problem of insufficient datasets, semi-supervised learning that utilizes the output results of deep learning models for unknown images will study. These methods will be a subject for future study.

REFERENCES

[1] *General Wastewater Standards, Ministry of the Environment*, https://www.env.go.jp/water/impure/haisui.html, Accessed in Nov. 2021 (in Japanese).

[2] *Industrial Wastewater Management in Japan*, https://www.env.go.jp/en/focus/docs/files/20120801-51.eps, 2012.

[3] M. B. Khan, H. Nisar and C. A. Ng, Image processing and analysis of phase-contrast microscopic images of activated sludge to monitor the wastewater treatment plants, *IEEE Access*, vol.6, pp.1778-1791, 2018.

[4] S. Selvan, J F. Joseph, K. D. Raj and B. J. Selvan, Analysis and classification of active sludge in wastewater, *Journal of Physics: Conference Series*, DOI: 10.1088/1742-6596/1916/1/012016, 2021.

[5] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press, http://www.deeplearningbook.org, 2016.

[6] C. Shorten and T. M. Khoshgoftaar, A survey on image data augmentation for deep learning, *Journal of Big Data*, vol.6, DOI: 10.1186/s40537-019-0197-0, 2019.

[7] C. Rachid, *Chemicals in the Food Industry: Toxicological Concerns and Safe Use*, Springer International Publishing, 2020.

[8] VoTT (Visual Object Tagging Tool), *Open Source Projects and Samples from Microsoft*, https://github.com/microsoft/VoTT, Accessed in Nov. 2021.

[9] C. Szegedy et al., Going deeper with convolutions, *arXiv.org*, arXiv: 1409.4842, 2014.

[10] K. He, X. Zhang, S. Ren and J. Sun, Deep residual learning for image recognition, *arXiv.org*, arXiv: 1512.03385, 2015.

[11] C. Szegedy, S. Ioffe, V. Vanhoucke and A. Alemi, Inception-v4, Inception-ResNet and the impact of residual connections on learning, *arXiv.org*, arXiv: 1602.07261, 2016.

[12] F. Chollet et al., *Keras*, https://keras.io, Accessed in Nov. 2021.

[13] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You Only Look Once: Unified, real-time object detection, *arXiv.org*, arXiv: 1506.02640, 2015.

[14] J. Redmon and A. Farhadi, YOLOv3: An incremental improvement, *arXiv.org*, arXiv: 1804.02767, 2018.