2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

Discrimination Method of Rebar and Concrete Lumps Using Gabor Filter for Automated Construction Demolition

Takahiro Ikeda[1], Shun Ito[1], Satoshi Ueki[1] and Hironao Yamada[1]

***Abstract*— To automate the process of breaking up a building into small pieces, it is necessary to discriminate rebars from concrete lumps to remove rebars. This paper proposes a method for discriminating rebar from concrete lumps using Gabor filters. Our discrimination system uses a filter bank consisting of multiple Gabor filters. Using the created filter bank, feature extraction is performed from each segmented region, and the image is classified into rebar and concrete chunks using a support vector machine. The system was run on images of actual rebar and concrete lumps and achieved a discrimination accuracy of 89.4%.**

I. INTRODUCTION

In recent years, the shortage of human resources and longer working hours at construction sites have become serious problems due to the decrease in the number of construction companies and construction workers and the aging of construction workers. In addition, the introduction of automation and robot technology has not progressed sufficiently in the construction industry due to high initial investment, the complexity of introducing automated processes, the segregation of tasks, and the lack of construction-specific tools, indicating that labor productivity is lower than in other industries [1]. Therefore, in recent years, the development of automation and robot technology in the construction industry has been desired to improve productivity, solve human resource shortages, and reduce costs.

This study focused on the removal of reinforcing steel bars that occur during the breaking up of concrete lumps at building demolition sites. Automation of rebar removal requires recognition of the rebar as a string-like object. Sato et al. and Fan et al. used Canny's filter [2] to detect parallel edges and enable recognition of string-like objects [3], [4]. Recognition using parallelism of edges is considered to be effective for recognizing rebar. However, the surface of lumps is highly uneven, and many unnecessary edges are detected. In addition, parallel lines may be detected from the lumps because the lumps are an artifact. Therefore, it is difficult to detect rebar from edge detection. In addition, there are studies that use three-dimensional data acquired from 3D cameras, etc., instead of two-dimensional camera images to recognize string-like objects. Tomoya et al. proposed an algorithm to recognize a string and its shape by using a point chain model for a 3D point cloud of the string acquired from a 3D camera [5]. Kajioka et al. proposed a method to discriminate strings from an environment in which strings and other objects are mixed together using 3D point cloud data [6]. In these studies, a 3D point cloud was acquired by a 3D camera and analyzed to recognize string-like objects. However, there is a problem that highly accurate 3D data cannot be obtained when the distance from the 3D camera is far or the object to be recognized is relatively small.

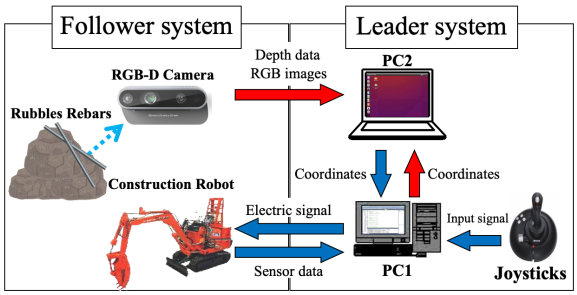


Fig. 1. Hardware overview.

Therefore, the purpose of this study is to develop a system that can recognize rebar and debris separately from a mixed environment of rebar and debris, and automatically remove the rebar. A filter bank consisting of multiple Gabor filters was used for the recognition of rebar and concrete lumps. Using the created filter bank, the Simple Linear Iterative Clustering (SLIC) algorithm was used to extract features from the segmented small regions, and a support vector machine was used for image classification. The recognition accuracy of the rebar and concrete lumps was evaluated using images of the actual rebar and lumps and the created model of the rebar and lumps. The effectiveness of the proposed system was confirmed by performing rebar removal using the rebar model and concrete model with an autopiloted construction machine. Figure 1 shows an overview of the hardware for this study. The detail of the hardware is described in [7].

II. DISCRIMINANT METHOD

This study developed a recognition system for rebar and concrete lumps, which consists of region segmentation, feature generation, image classification, and rebar location estimation. The images of rebar and concrete lumps were actual images collected from an internet image search and model images made of styrofoam. Fig. 2 shows the real images created for the example, and Fig. 3 shows the model images.

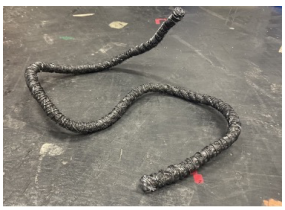
*This work was not supported by any organization

[1]All authors are with Faculty of Engineering, Gifu University, 1-1, Yanagido, Gifu, 501-1193, Japan. Ikeda is a corresponding author. ikeda.takahiro.w5@f.gifu-u.ac.jp, ito.shun.3.nov@gmail.com, {ueki.satoshi.m8, yamada.hironao.a5}@f.gifu-u.ac.jp

(a) Rebar images

(b) Concrete lump images

Fig. 2. Sample images of real rebar and rubble.



(a) Rebar model

(b) Concrete lump model

Fig. 3. Rebar model and rubble model.

This method discriminates between rebar and concrete lumps by the following process. (1) Acquire RGB images of the rubble. (2) The image is divided into superpixels by SLIC. (3) Apply a Gabor filter to the superpixels, and (4) extract the features of the superpixels using BoVW. (5) Classify the superpixels by SVM based on the features and label the superpixels. The objective of this research is to automatically remove rebar from concrete lumps. The rebar pixel closest to the center of gravity of the pixels determined to be a rebar was used as the target tip position of the construction robot.

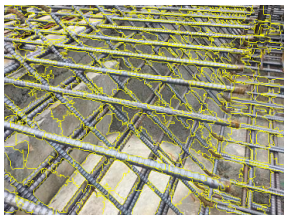
A. Region segmentation by SLIC

The superpixel algorithm [8] is often used in region segmentation for preprocessing of object recognition. This study uses the SLIC algorithm [8], [10], [11], [12] for region segmentation. The collected images were used to experimentally determine the values that successfully separate the rebar and the concrete lumps in a relatively large region. The results show that the initialization size of the region is 35 pixels and the ratio coefficient between color and position is 0.06. The images before and after adaptation of the SLIC algorithm are shown in Fig. 4(a) and Fig. 4(b), respectively. In this study, superpixels were treated as a single input image, and object recognition was performed by classifying each superpixel. Since the shape of a superpixel is complex, when each superpixel is extracted by rectangle, the extracted superpixel and its surrounding area are included as shown in Fig. 5(a). Therefore, each superpixel area was extracted as an image masked with black color except for the extracted superpixels as shown in Fig. 5(b).

B. Feature extraction using Gabor Filter

This section describes methods for extracting image features and quantizing feature vectors. This study used a Gabor filter [13] to extract image features. The Gabor function consists of a Gaussian function and sine and cosine functions and can extract specific spatial frequencies for each direction in a local region around each point in the image. The Gabor

(a) Input image.



(b) Output image.

Fig. 4. Region segmentation.

(a) Superpixel with surrounding area.

(b) Processed superpixel with surrounding area.

Fig. 5. Process superpixels into a rectangular image.

function $g(x, y, \sigma, \lambda, \gamma, \theta)$. is defined as follows;

$$\begin{aligned} g &= \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(\frac{2\pi x'^2}{\lambda}\right), \quad (1)\\ x' &= x\cos\theta + y\sin\theta,\\ y' &= -x\sin\theta + y\cos\theta, \end{aligned}$$

where, θ is the angle of the filter, σ is the standard deviation, λ is the wavelength, and γ is the spatial aspect ratio. Each parameter should be given an optimal value depending on the input image. Previous studies using Gabor filters generally employ 8-, 12-, and 16-angle directions. 12- and 16-angle directions can extract a large number of features but have the disadvantage of increasing redundancy. In this study, a Gabor filter was created with 8 directions, because 8 directions can extract enough features for oblique components.

Since the Gabor filter extracts specific frequencies of an image as features, the setting of the parameters λ and σ is very important. According to the literature [14], it is concluded that the optimal combination of λ and σ is the one according to following equations:

$$\sigma = 0.5\lambda, \quad and \quad (2)$$

$$\sigma = 0.7\lambda. \quad (3)$$

In addition, λ was set to

$$\lambda = \sqrt{2}, \quad 2\sqrt{2}, \quad 3\sqrt{2} \quad (4)$$

with reference to the candidate points in the literature [14]. σ was set values satisfying equations 2 and 3 were used. The parameter γ was set to $\gamma = 0.5$ based on the candidate points in the literature [15].

Image features were extracted by computing the convolution of superpixel images using a Gabor filter. Pixel

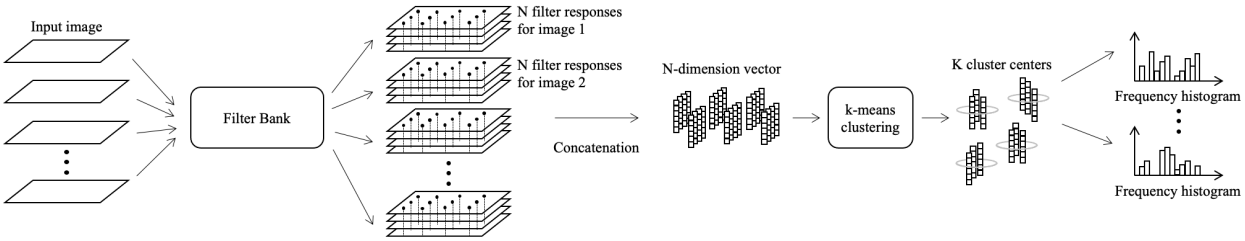


Fig. 6. Processing image with filter bank and BoVW

$P_{out}(x, y)$ in the output image is defined as follows:

$$P_{out} = \sum_{i=0}^{N} \sum_{j=0}^{N} P_{in}(x+i, y+j) \times f(i, j), \qquad (5)$$

where, $P_{in}(x, y)$ is the input image pixel, and $f(i, j)$ is the $N \times N$ size kernel value. Note that in this study, the kernel size was set to 21×21 because it is the smallest range in which the wave portion of the Gabor filter can fit. Parameters θ, γ, and N were fixed, and λ and σ are varied to create a Gabor filter. 8-direction Gabor filters can be created by uniquely selecting λ and σ. This is called a filter bank. The filter bank consists of a collection of filters, which are used to divide an image into several spatial frequency bands. Therefore, this study created eight filter banks, including six filter banks that could be created when λ was uniquely selected, and two filter banks that could be created when all three types of λ were combined.

In convolutional computation, a rectangular image filled with black color except for the region of interest is used as the input image. Therefore, if convolution is simply performed on all pixel regions, useful features cannot be obtained. To solve this problem, image completion of the input image was performed. In research on image completion processing, completion is often performed using two types of approaches: smooth completion of pixel values [16] and texture composition-based methods [17], [18]. In the approach that smoothly completes pixel values, completion is performed by smoothly propagating pixel values from the boundary of the missing region inward. In the approach of texture composition, a non-deficient region is copied and pasted to the missing area. In this study, the method that smoothly completes pixel values is inappropriate because the missing areas are located at the edges of the input image. Therefore, an approach based on texture synthesis was used for image completion. In the texture composition-based approach, it is necessary to select the source region for copying, taking into account the natural appearance of the missing region. However, the superpixels handled by this system are segmented into regions with the same properties, eliminating the need to select a source region. Therefore, image completion was performed by extracting the largest rectangular region from the region of interest in the input image and attaching it to the missing region.

In classification such as object recognition, it is difficult to use the extracted features without special processing. In this study, an algorithm called Bag of visual words (BoVW) [19] was used to quantize feature vectors. This BoVW algorithm considers an image as a document and local features in the image as words and identifies categories by using histograms of local features as features of the image. The flow up to the creation of feature vectors is shown below (Fig. 6).

1) Obtain filter responses from all input images with image completion using a filter bank consisting of N Gabor filters.
2) Concatenate the N filter responses obtained in step 1, including the RGB values if RGB values are added to the features.
3) The elements (i, j) of each filter response are extracted in a concatenated state, and this is used as the feature vector at pixel (i, j). The feature vectors are extracted only from the range of the input image that is not missing before image completion.
4) The k-means method [20] is applied to the extracted feature vectors to divide the N-dimensional feature vector space into k clusters.
5) Records how many clusters the feature vectors extracted from each image belong to in the N-dimensional space.
6) A histogram is created based on the recorded results, and this is used as the new feature vector.

In this study, support vector machine (SVM, [9]) is employed to classify two classes: rebar and concrete lump. The kernel function is an RBF kernel, and the cost parameter C and the hyperparameter γ of the RBF kernel are automatically optimized by grid search. The hyperparameters C and γ were determined experimentally. Table I shows the candidate values of C and γ.

TABLE I
CANDIDATE VALUES AND TYPES OF HYPERPARAMETERS OF SVM

Kernel	RBF
C	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
γ	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0

III. EXPERIMENT

A. Experimental setup

This section describes an experiment to evaluate the accuracy of the proposed classification algorithm. There are

various types of reinforcing bars and debris present at building demolition sites in terms of shape and color. Moreover, because the demolition work is performed outdoors, the working environment is not constant due to the effects of sunlight and other factors. Therefore, the rebar and gable recognition system must be robust to the effects of shape, color, and sunlight.

This accuracy evaluation used real images and model images of rebars and concrete lumps. The model images were collected with the rebar and concrete lump models randomly placed within 0.5 m to 1.5 m from the RGB-D camera. The SLIC algorithm was applied to all these images. The generated superpixel images were labeled as correct for all images: 8709 real rebar images, 8709 real concrete lump images, 2376 rebar model images, and 2625 concrete lump model images. 75% of the superpixel images database were used as training data and 25% of that were used as test data to evaluate the accuracy of the discrimination system. Figure 7 shows the superpixel images of real rebars and concrete lumps, and Fig. 8 shows the superpixel images of them. To evaluate the effect on the accuracy of the system from the sunlight, the experiment used the shade image where the model was in the shade and the sun image where the model was in the sun. To evaluate the robustness of the model under a sunlight environment, 1000 sun images were used as training images, and 250 sun images were used as test images (sun model). Also, 1000 shade images were used as training images and 250 sun images were used as test images (shade model). The examples of superpixel images used in the accuracy evaluation are shown in Fig. 9 for the model images in the sun and in Fig. 10 for the model images in the shade.

Several filter banks consisting of Gabor filters with different parameters were prepared, and the discrimination accuracy was evaluated for each of them. The parameter choices for the Gabor filters are as follows:

- wavelength: $\lambda = \sqrt{2}, 2\sqrt{2}, 3\sqrt{2}$,
- standard deviation: $\sigma = 0.5\lambda$, $\sigma = 0.7\lambda$,
- angle: $\theta = \frac{\pi}{8}, \frac{2\pi}{8}, \frac{3\pi}{8}, \frac{4\pi}{8}, \frac{5\pi}{8}, \frac{6\pi}{8}, \frac{7\pi}{8}$,
- spatial aspect ratio: $\gamma = 0.5$,
- kernel size: $N = 21$.

The parameters θ, γ, and N were fixed and applied to all Gabor filters created. Six filter banks were created when λ was uniquely selected, and two filter banks were created when all three λ parameters were combined. The experiment also evaluated the accuracy of adding the RGB value of each pixel to the features.

The discrimination accuracy A is defined as follows:

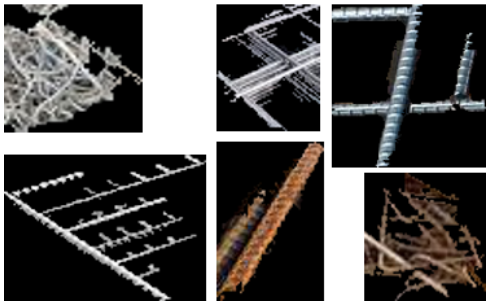
$$A = \frac{T}{n}, \tag{6}$$

where n is the number of all test images and T is the number of images for which the predicted class matches the true class.

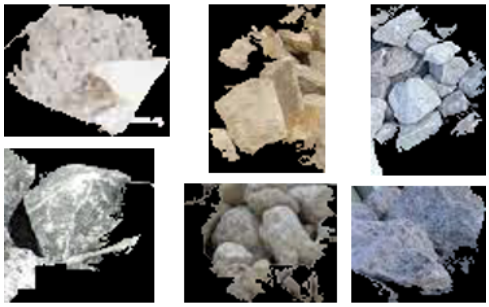
B. Experimental result

Table II shows the accuracy of the discriminator when discriminating between images of actual rebar and concrete lumps for the Gabor filter parameter configurations described above. According to Table II, the relationship with $\sigma = 0.5\lambda$ is more accurate than the relationship with $\sigma = 0.7\lambda$. Since the σ parameter expresses the standard deviation of the Gaussian function, the smaller σ is, the more elongated the Gaussian function is. A Gabor filter with $\sigma = 0.5\lambda$ is more undulating than a filter with $\sigma = 0.7\lambda$, and therefore more strongly extracts specific frequencies in the image. Therefore, the discriminator is considered to be more accurate when $\sigma = 0.5\lambda$. Since λ extracts frequencies contained in the image, the accuracy varies depending on the image being handled. For the images used in this study, the discriminator was highly accurate when $\lambda = 2\sqrt{2}$. The accuracy was also higher when three scales were combined, such as $\lambda = \sqrt{2}, 2\sqrt{2}, 3\sqrt{2}$. For the accuracy validation, rebar images at various scales were collected. Therefore, it is considered that a filter bank combining multiple λs, from which features can be extracted at multiple scales, would have resulted in higher accuracy.

Table III shows the accuracy of the discriminator when the images of the models of rebar and concrete lumps (indoor environment) are used for discrimination. The validation in Table III was conducted using model images for training and



(a) Images of real rebars.



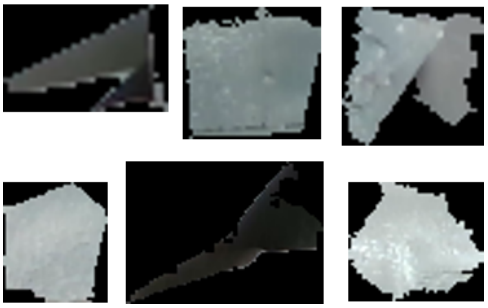
(b) Images of real concrete lumps.

Fig. 7. Superpixel image examples of real rebars and concrete lumps.

TABLE II
DISCRIMINATION ACCURACY FOR IMAGES OF REAL OBJECTS

λ	$\sigma = 0.5\lambda$	$\sigma = 0.7\lambda$
$\sqrt{2}$	0.870	0.846
$2\sqrt{2}$	0.888	0.850
$3\sqrt{2}$	0.873	0.828
$\sqrt{2}, 2\sqrt{2}, 3\sqrt{2}$	0.894	0.874

(a) Images of rebar models.



(b) Images of concrete models.

Fig. 8. Superpixel image examples of models of rebars and concrete lumps.

(a) Images of rebar models.

(b) Images of concrete models.

Fig. 9. Superpixel images of rebar and concrete lump models in sun.

TABLE III
DISCRIMINATION ACCURACY FOR IMAGES OF MODEL OBJECTS

λ	$\sigma = 0.5\lambda$	$\sigma = 0.7\lambda$
$\sqrt{2}$	0.998	0.996
$2\sqrt{2}$	0.994	0.989
$3\sqrt{2}$	0.993	0.986
$\sqrt{2}, 2\sqrt{2}, 3\sqrt{2}$	0.996	0.992

TABLE IV
DISCRIMINATION ACCURACY FOR IMAGES OF SUN MODEL

λ	$\sigma = 0.5\lambda$	$\sigma = 0.7\lambda$
$\sqrt{2}$	0.996	0.994
$2\sqrt{2}$	0.976	0.982
$3\sqrt{2}$	0.974	0.956
$\sqrt{2}, 2\sqrt{2}, 3\sqrt{2}$	0.984	0.984

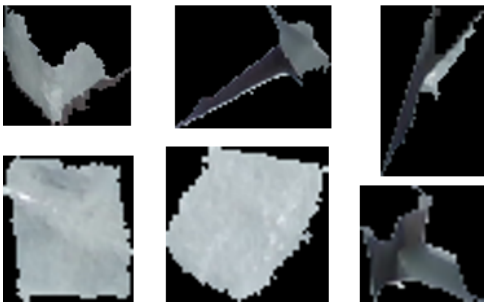
testing. Therefore, the classification was very accurate. The relationship between σ and λ was the same as for the real rebar and concrete images, with $\sigma = 0.5\lambda$ being the best. In the model image, the highest accuracy was obtained when $\lambda = \sqrt{2}$. The method that combines three scales has a lower accuracy than the filter bank with one scale. This is because the model images were collected from 0.5 m to 2.0 m from the camera, and there were fewer images with different scales than the real object images. In addition, the method of combining multiple scales increases the dimensionality of the feature vectors, which squeezes the memory and reduces the computation speed. Therefore, if the scale of the captured object is relatively constant from the camera's point of view, there is no need to combine multiple scales.

Table IV shows the accuracy of the discriminator when the images of the models in the sun are used for learning and evaluation (sun model). According to Table IV, the best results were obtained when $\sigma = 0.5\lambda$ and $\lambda = \sqrt{2}$, similar to the results obtained with the images of the model in the indoor environment, whose results are presented in Table 2. The results for multiple λ combinations were slightly inferior to those for the above best result. This is because the sun model was taken in the same distance between camera and model as when the model was taken indoors.

Table V shows the accuracy of the discriminator when the images of the models in the shade are used for learning and the images in the sun are used for the evaluation (shade model). According to Table V, the best results were obtained when $\lambda = \sqrt{2}$, but the results for $\sigma = 0.5\lambda$ and $\sigma = 0.7\lambda$ were equally accurate. The results for multiple λ combinations were slightly inferior to those for the result of $\lambda = \sqrt{2}$. This is because the shade model was taken in the same distance between camera and model as when the model was taken indoors. In addition, the shade model was slightly inferior to the sun model (Table IV) because the training and test data for this shade model used data from different shooting environments. However, the drop in discrimination accuracy was 2.0% for $\lambda = \sqrt{2}$, 1.0% for $\lambda = 2\sqrt{2}$, 1.6% for $\lambda = 3\sqrt{2}$, and 1.4% for multiple λ (when $\sigma = 0.5\lambda$). This means that our method is robust against sunlight.

Finally, it was confirmed that the automatic piloted construction equipment could recognize rebars by the proposed system and remove them (Fig. 11). The distance to the recognized rebar was measured using a depth camera. As a

(a) Images of rebar models.



(b) Images of concrete models.

Fig. 10. Superpixel images of rebar and concrete lump models in shade.

TABLE V

DISCRIMINATION ACCURACY FOR IMAGES OF SHADE MODEL

λ	$\sigma = 0.5\lambda$	$\sigma = 0.7\lambda$
$\sqrt{2}$	0.978	0.978
$2\sqrt{2}$	0.966	0.962
$3\sqrt{2}$	0.958	0.942
$\sqrt{2}, 2\sqrt{2}, 3\sqrt{2}$	0.970	0.968

result of the experiment, the construction machine was able to magnetically adhere to the rebar in all 30 trials.

IV. CONCLUSIONS

This paper described a method for discriminating between concrete lumps and rebars to automate the work of breaking small pieces of concrete by construction equipment. The discriminator divides the input image into superpixels using SLIC, extracts features from them using a Gabor filter, and discriminates rebar and concrete lumps using SVM. Experimental results showed that the discrimination accuracy was 89.4% for the real object images. In an experiment using a construction machine to automatically remove the rebar models, a 100% removal rate was achieved.

REFERENCES

[1] Construction Industry Handbook, Japan Construction Industry Association, 2021.

[2] J. Canny, "A Computational Approach to Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-8, no. 6, pp. 679-698, 1986.

[3] Kenji Sato, Kei Okada and Masayuki Inaba, "Humanoids' daily life support actions using image processing that recognizes string-like objects," Proc. of the 2006 JSME Conference on Roboics and Mechatronics, 2A1-D26, 2006.

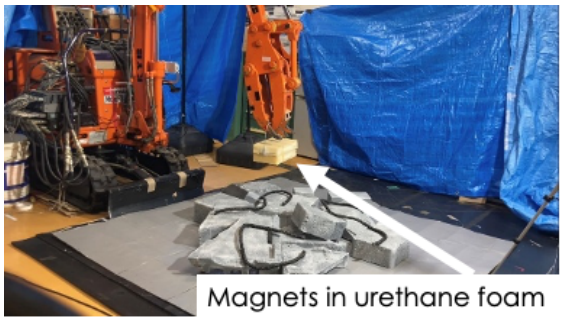


Fig. 11. Rebar removal experiment.

[4] Zheming Fan, Takeshi Ohashi and Toyohiro Hayashi, "Research on cord-like object recognition system," Information Processing Society of Japan Signal Technical Report, pp. 1-8, 2019.

[5] Tomoya Shirakawa, Keisuke Mukai, Takayuki Matsuno, Akira Yanou and Mamoru Minami, "Form Modeling of a String And Recognition Using Distance Sensor," Proc. of the 2015 JSME Conference on Robotics and Mechatronics, 2A1-R01, 2015.

[6] Daisuke Kajioka, Hiromasa Onaka, Zhang Chi, Yuichiro Toda, Takayuki Matuno and Mamoru Minami, "Recognition of linear deformable objects using point cloud data for various objects mixedly placed," Proc. of the 2021 JSME Conference on Robotics and Mechatronics, 2P2-H09, 2021.

[7] Takahiro Ikeda, Satoshi Ueki, Kazuma Shinkai and Hironao Yamada, "Depth Distribution Split Labeling for Rubble Recognition of Crushing Machine," Proc. of 2022 IEEE International Conference on Robotics and Automation, pp. 8150-8156, 2022.

[8] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk, "SLIC Superpixels Compared to State-of-the-art Superpixel Methods," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 11, pp. 2274-2282, 2012.

[9] B. Boser, I. Guyon and V. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," Proc. of the 5th Annual Workshop on Computational Learning Theory (COLT'92) , pp. 144-152, 1992.

[10] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient Graph-Based Image Segmentation," International Journal of Computer Vision, Vol. 59, No. 2, pp. 167-181, 2004.

[11] A. Vedaldi and S. Soatto, "Quick Shift and Kernel Methods for Mode Seeking," Proc. of European Conference on Computer Vision, pp. 705-718, 2008.

[12] P. Neubert and P. Protzel, "Compact Watershed and Preemptive SLIC: On Improving Trade-offs of Superpixel Segmentation Algorithms," International Conference on Pattern Recognition, pp. 996-1001, 2014.

[13] D. Gabor, "Theory of communication," Journal of the Institution of Electrical Engineer, Vol. 93, pp. 429-441, 1946.

[14] Yoshihiko Hamamoto, Shunji Uchimura, Masamori Waranabe, Tetsuya Yasuda, Yoshihiro Mitami and Shingo Tomita, "A gabor filter-based method for recognizing handwritten numerals," Pattern Recognition, Vol. 31, No. 4, pp. 395-400, 1998.

[15] Francesco Bianconia, Antonio Fernández, "Evaluation of the effects of Gabor filter parameters on texture classification," Pattern Recognition, Vol. 40, pp. 3325-3335, 2007.

[16] M. Bertalmio, A.L. Bertozzi and G. Sapiro, "Navier-Srokes, Fluid Dynamics, and Image and Video Inpainting," Proc. IEEE Conf. Computer Vision and Pattern Recognition, Vol. 1, pp. 355-362, 2001.

[17] Y. Wexler, E. Shechtman and M. Irani, "Space Time Video Completion," Proc. IEEE Int. Conf. on Pattern Recognition, pp.120-127, 2004.

[18] A. Efros and T. Leung, "Texture Synthesis by Non-Parametric Sampling," Proc. Int. Conf.Computer Vision, pp. 1033-1038, 1999.

[19] Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski and Cédric Bray, "Visual categorization with bags of keypoints," European Conference on Computer Vision (ECCV2004) Workshop on Statistical Learning in Computer Vision, pp. 59-74, 2004.

[20] James MacQueen, "Some methods for classification and analysis of multivariate observations," Proc. of the fifth Berkeley symposium on mathematical statistics and probability, Vol. 1, No. 14, 1967.