

DETECTION AND MONITORING OF BEACH LITTER USING UAV IMAGE AND DEEP NEURAL NETWORK

S.H. Bak¹, D.H. Hwang¹, H.M. Kim¹, H.J. Yoon^{1,*}

¹ Pukyong National University, Division of Earth Environmental System Science, 48513 Nam-gu Busan, South Korea -
yoonhj@pknu.ac.kr

KEY WORDS: Marine Debris, Unmanned Aerial Vehicle, Neural Network, Deep Learning, Marine Pollution

ABSTRACT:

Beach litter destroys marine ecosystems and creates aesthetic discomfort that lowers the value of the beach. In order to solve this beach litter problem, it is necessary to study the generation and distribution pattern of waste and the cause of the inflow. However, the data for the study are only sample data collected in some areas of the beach. Also, most of the data covers only the total amount of beach litter. UAV(Unmanned Aerial Vehicle) and Deep Neural Network can be effectively used to detect and monitor beach litter. Using UAV, it is possible to easily survey the entire beach. The Deep Neural Network can also identify the type of coastal litter. Therefore, using UAV and Deep Neural Network, it is possible to acquire spatial information by type of beach litter.

This paper proposes a Beach litter detection algorithm based on UAV and Deep Neural Network and a Beach litter monitoring process using it. It also offers optimal shooting altitude and film duplication to detect small beach litter such as plastic bottles and styrofoam pieces found on the beach.

In this study, DJI Mavic 2 Pro was used. The camera on the UAV is a 1-inch CMOS with a resolution of 20MP. The images obtained through UAV are produced as orthoimages and input into a pre-trained neural network algorithm. The Deep Neural Network used for Beach litter detection removed the Fully Connected Layer from the Convolutional Neural Network for semantic segmentation.

1. INTRODUCTION

1.1 General Instructions

Marine litter is classified as coastal litter, marine litter, and immersed litter. It is known that about 70% or more of these are generated by various municipal wastes and industrial wastes, and they are introduced into rivers through land(Chung *et al.*, 2012).

Among these, coastal litter waste destroys marine ecosystems and creates a sense of aesthetic discomfort that hampers the utility value of the beach. In order to solve the problem of shore-based garbage collection, it is necessary to study the generation and distribution pattern of garbage and the cause of the influx. However, the data for the study are only available for sampling in some areas of the beach due to time and economic conditions. In the case of marine litter monitoring in Korea, a sampling interval of 100 m in length and 20 m in width is set to collect samples and generalize the whole area(Jang *et al.*, 2011). However, this monitoring method has a considerable error factor(Park *et al.*, 2005). Therefore, it is necessary to investigate the amount of coastal catching garbage generated on the whole coast, and to establish a new monitoring method to overcome the limit of cost and time.

In recent years, remote sensing based monitoring techniques have been actively researched to solve these problems. Kako *et al.*(2010) calculated the amount of coastal litter waste by applying the threshold to images acquired using a webcam, and Kakashima *et al.*(2011) used Balloon Aerial Photography. In addition, Jang *et al.*(2011) calculated the generation rate by applying the threshold to the image of the unmanned aerial vehicle, and Bao *et al.*(2018) applied the two-step threshold filtering method to the image of the unmanned airplane.

The remote sensing based monitoring method has an advantage that it can estimate the amount of occurrence in a short time at a

low cost. However, the thresholding technique is effective when the variance among classes to classify is large, but there is a possibility of misclassification if not. It is therefore difficult to detect objects with similar colors to the background sand or gravel.

In this study, we propose a Deep Neural Network based coastal garbage detection algorithm to improve the disadvantage of the thresholding method.

2. DATA AND METHOD

2.1 Study area

The study area is Heungnam beach(Fig. 1) in Geoje, located on the South Sea of Korea, which falls in the marine climate zone, with an annual mean temperature of 25°C and annual accumulated precipitation of more than 2000mm.

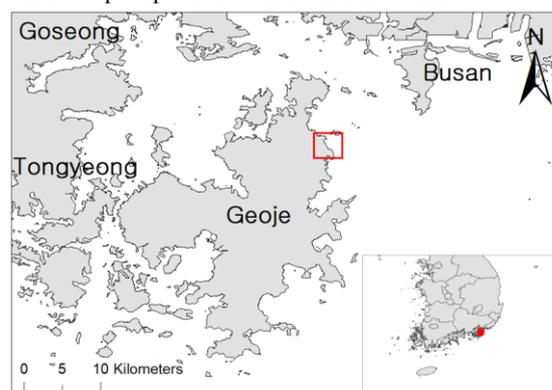


Figure 1. Study area(Heungnam beach, red box) and location of the study area in the South Korea(red dot)

2.2 UAV system platform

Unmanned aircraft are divided into multi-rotor and fixed wing according to the flight type. Fixed wing unmanned airplanes fly on wings in the form of a regular aircraft with fixed wings on the fuselage. Therefore, it is possible to fly for a long time with a small power by using a small number of propellers. On the other hand, multi-rotor has four or more propellers. Therefore, relatively high power is required and flight time is short. However, since it is possible to take off and land vertically, it does not depend on the place.

The study area, Heungnam beach, is surrounded by mountains and the width of the beach is narrow, making it difficult to operate fixed-wing unmanned aircraft. Therefore, this research utilized DJI's multi-rotor type unmanned aerial vehicle, MAVIC 2 PRO. The MAVIC 2 PRO is equipped with a 20MP (4,864 × 3,648 pixels) RGB camera and the flight time is approximately 25 minutes.

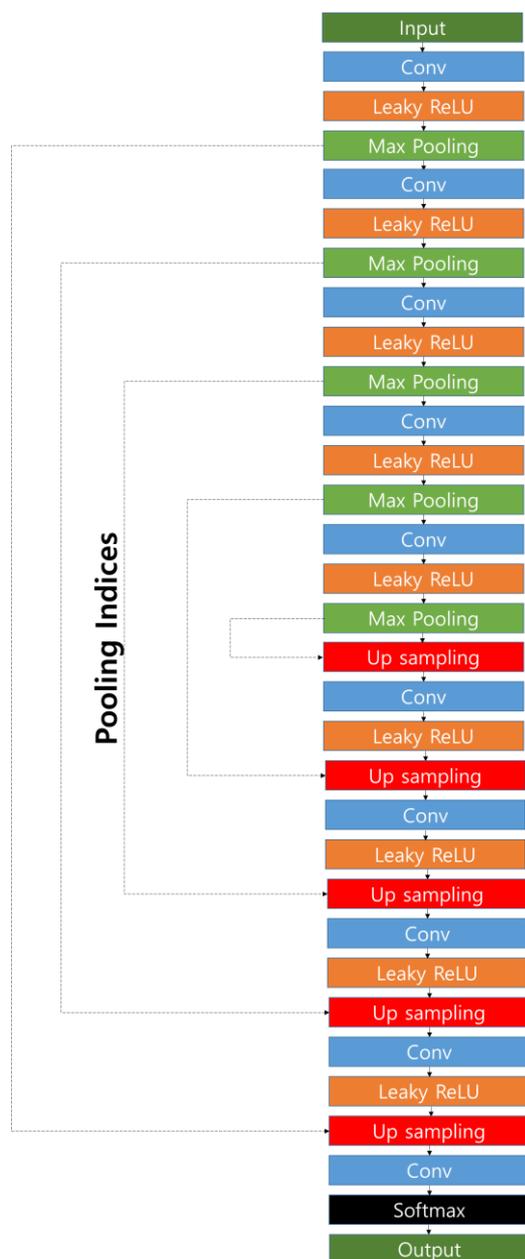


Figure 2. Structure of Neural Network

2.3 Deep Neural Network for detecting beach litter

In this study, we used the structure of SegNet (Badrinarayanan *et al.*, 2017) developed by the University of Cambridge to detect Beach litter.

SegNet is a Deep Neural Network for semantic segmentation and can be classified in pixel units. SegNet eliminates the full connection layer of VGG-16 and is made in the form of Encoder-Decoder. Encoder performs Feature extraction through Convolution and Max Pooling, and Decoder performs Classification by Pixel unit through Up Sampling and Convolution. In this case, Up Sampling is a process to reinforce information lost by Max Pooling in the Encoder part. It stores information of the biggest value extracted when Max Pooling is performed in each pooling layer of Encoder, and it is used as Pooling index.

The neural network used in this study consists of 5 layers of encoders and decoders respectively, Leaky ReLU (Rectifier Linear Unit) for the activation function and Softmax for classification in the last layer of the decoder (Fig. 2).

2.4 Process of detecting beach litter

The detection process of the beach litter using UAV video and Deep Neural Network is composed of 3 steps in total (Fig. 3).

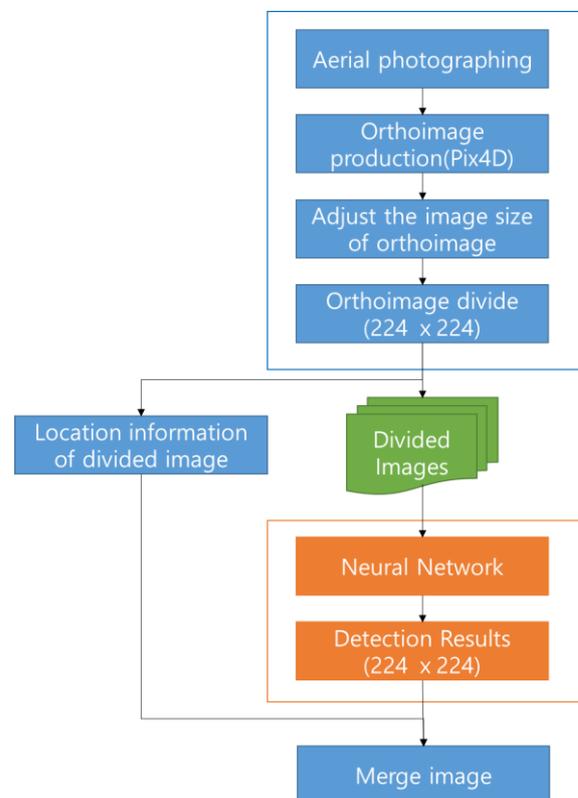


Figure 3. Process of detecting beach litter

1. Image acquisition and preprocessing.
2. Detection using Neural Network.
3. Postprocessing

In step 1 (Image acquisition and preprocessing), images are taken using UAV and orthoimages are produced. At this time, the shooting altitude was set to 15 meters in consideration of the size of the beach litter and the ground sampling distance (GSD), and the end overlay and the lateral overlay were set at 85%.

respectively. UAV images were produced by orthoimage using Pix4D. The orthoimage is divided into 224×224 pixel size image for input to the neural network, and the position information of each divided image is recorded.

In step 2 (Detection using Neural Network), the divided images generated in step 1 are sequentially input to the Neural Network to detect beach litter.

The UAV image used in this study is imbalance data that is captured at the time when the amount of the beach litter is small and the ratio of the pixel occupied by the detected object in the image is less than 5%. Most pixels correspond to Background, and only some pixels are detected (Styrofoam, Plastic, Glass, etc.).

Imbalanced data means that the number of labels for a particular class is significantly smaller or greater than for other classes, which can seriously degrade the performance of machine learning-based models (Bak *et al.*, 2018). In general, resampling resolves the unbalanced data problem. The resampling method includes an oversampling technique that amplifies the number of classes with fewer labels and undersampling that uses only some of the data of a class with a large number of labels.

In the case of undersampling, a class with a large number of labels can cause performance degradation if it is a phenomenon with various examples (Bak *et al.*, 2018). In this study, we used Oversampling because Background class, which has a large number of labels, is a class with various examples (gravel, dry sand, wet sand, etc.).

In step 3 (Postprocessing), the detection result image generated in step 2 is merged and restored to orthoimage size using the position information of each divided image recorded in step 1.

3. RESULT AND DISCUSSION

3.1 Training results of Deep Neural Network.

The Deep Neural Network used in this study was learned 1,000 times and data augmentation was done every epoch to solve the imbalanced data problem. Data augmentation was performed by randomly cropping and rotating or reversing the image from the image containing the detection object. For each Epoch, 6 batch images were generated and used for learning. The ratio of the image to be detected to the background image was set to 1:2. The accuracy of the learning result model was about 98.2% (Fig. 4).

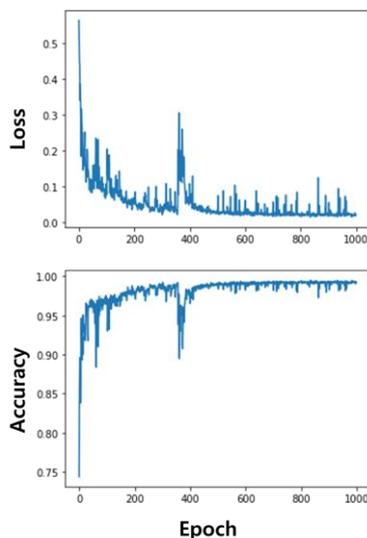


Figure 4. Loss and accuracy for iteration

3.2 Detection Result

The results of detection of Styrofoam in beach litter against test image are shown in Fig. 5. The detection target Styrofoam was well detected, and some noises existed, but it was found that they are generally well detected without confusion with other beach litters.

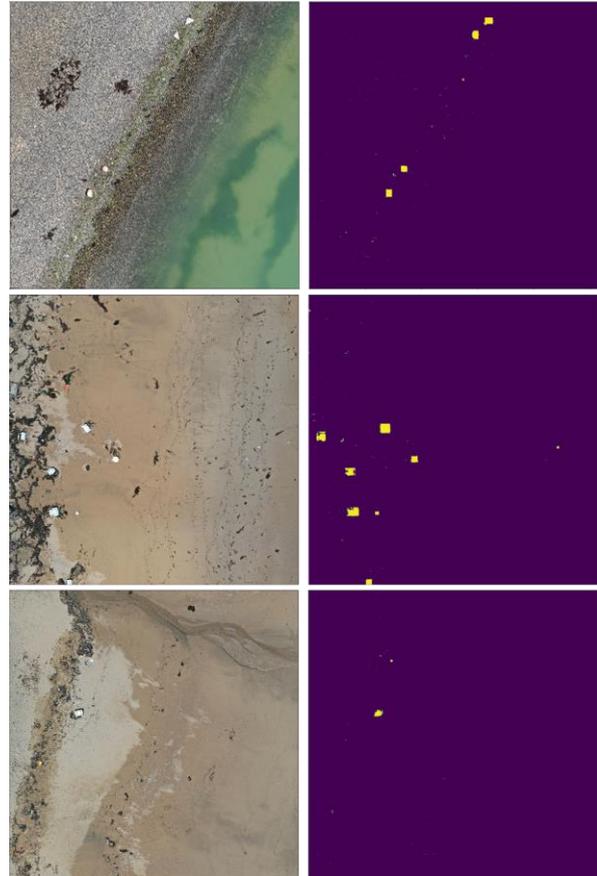


Figure 5. Styrofoam detection result from deep neural network (left : input image, right : detection result)

Styrofoam is a beach litter that is susceptible to breakage due to poor durability of materials and exists in various forms in the marine environments. Also, since most of the styrofoam products are white, the surface is likely to be contaminated during floatation at sea level. Therefore, styrofoam has various types of discoloration, which can cause difficulties in detection using color or reflectivity. Nevertheless, the Deep Neural Network was able to detect Styrofoam well. In particular, it was confirmed that even if the surface is contaminated and discolored, the detection is possible (Fig. 6)

The reason for this result is that the Convolution Neural Network has the ability to extract higher dimensional features in addition to the apparent appearance and color. Convolution Neural Network can utilize both Local Feature and Global Feature through Convolution Layer and Pooling Layer. In addition, it is possible to utilize high-dimensional features by extracting local feature and global feature from the global feature extracted from the previous layer through several layers of Convolution Layer and Pooling Layer (Kamishima *et al.*, 2015).

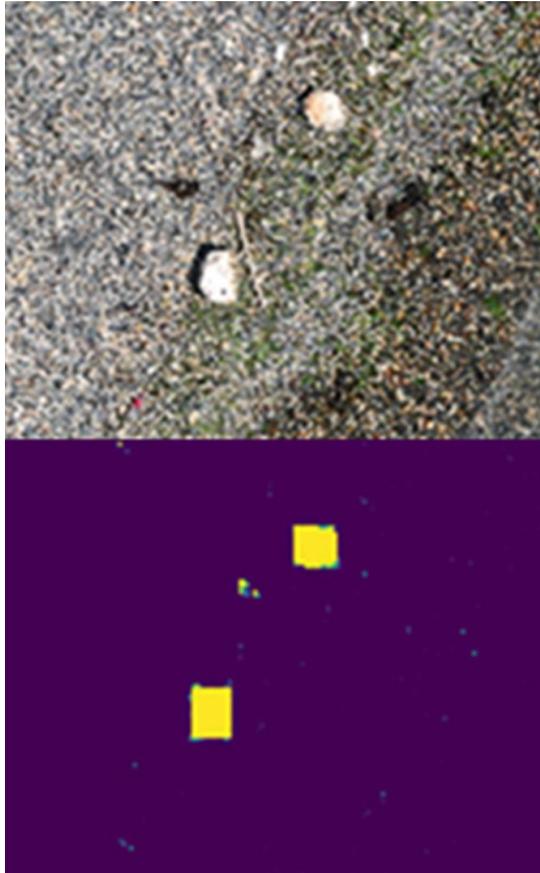


Figure 6. Detection result of styrofoam contaminated surfaces(top : input image, bottom : detection result)

In the previous studies, only the color or reflectivity is used in the process of extracting the Global Feature from the input image. However, in the case of CNN-based algorithms, it is possible to repeat local and global feature extraction by alternately stacking Convolution Layer and Pooling Layer. Unlike previous studies, color information as well as morphological information can be used as a feature. This makes it possible to distinguish objects that have the same color even if they have different pattern patterns.

4. CONCLUSION

In this study, we attempted to detect beach litter in unbalanced data environment using UAV image and Convolutional Encoder-Decoder. We oversampled the Beach litter class to solve the unbalanced data problem. Oversampling is implemented by data augmentation which rotates or reverses the original image. As a result, the Convolution Encoder-Decoder showed good performance in detecting irregular shapes such as Styrofoam and targets with various colors. Previous studies also have the ability to distinguish a particular class within the beach litter, further from the distinction between Background and beach litter.

These results were obtained because of the feature extraction method of CNN - based neural networks. CNN has a special structure called Convolution Layer and Pooling Layer to extract various features. In addition, when multiple layers of Convolution Layer and Pooling Layer are used, it is possible to extract the Global Feature again from the Global Feature extracted in the previous process.

The characteristics of CNN can be useful for identification of beach litter where various pattern shapes exist due to breakage and corrosion.

REFERENCES

- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017, SegNet : A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481-2495.
- Chung, I.H., Park, S.W., Kim, M.C., 2012, Generation Quantity and the Physicochemical Properties of Marine Litter Occurred in island area, *J. of Korean Society of Environmental Technology*, 13(4), 293-300.
- Jang, S.W., Lee, S.K., Oh, S.Y., Kim, D.H., Yoon, H.J., 2011, The Application of Unmanned Aerial Photography for Effective Monitoring of Marine Debris, *J. of the Korean Society of Marine Environment & Safety*, 17(4), 307-314.
- Kamishima, T., Asoh, H., Yasuda, M., Maeda, S., Okanohara, D., Okatani, T., Kubo, Y., Bollegala, D., 2015, *Shinsou gakushu – Deep Learning*. Kindaikagakusha, Tokyo.
- Park, S.S., Kang, H.Y., 2005, The Quantity and Characteristics of Marine Debris Collected from the Coastline in Jeonnam, *J. Korea Society of Waste Management*, 22(2), 203-212.