International Journal of Advanced and Applied Sciences, 11(3) 2024, Pages: 36-45

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Developing a plogging activity tracking app using deep learning for image recognition

Jung-Been Lee*, Taek Lee, Jeong-Dong Kim, In-Hye Yoo, Da-Bin Kim, Jung-Yeon Park

Division of Computer Science and Engineering, Sun Moon University, Asan, South Korea

ARTICLE INFO

Article history:
Received 1 September 2023
Received in revised form
3 January 2024
Accepted 6 February 2024

Keywords:
Deep learning
Object classification
Mobile application
Trash management
Plogging

ABSTRACT

Plogging is an activity that combines jogging with picking up litter, and participants often share their efforts on social media. However, the repetitive bending involved in plogging may cause back strain, and manually entering details such as the location and quantity of litter could slow the spread of this activity. This study sought to create and test a deep learning application to automatically monitor and record plogging by identifying the type and quantity of litter. We employed Convolutional Neural Networks (CNN) and YOLOv5 to develop an image recognition model. This model allowed users to easily log their plogging efforts by simply taking a photograph, removing the need to manually input the litter details. Moreover, we proposed a reward system that uses the collected trash amount and the distance covered to promote competition among users. We developed the first application that uses deep learning to automatically identify litter for tracking plogging activities. However, as this application was only a prototype, no comparative studies or usability tests were done. In future research, we plan to assess the application's usability and compare it with other similar applications.

1. Introduction

Garbage originating from terrestrial sources often contaminates rivers, subsequently leading to significant trash dispersion in coastal regions. Instances of natural disasters, such as heavy rainfall or monsoons, exacerbate this contamination by discharging large quantities of previously land-confined domestic trash into the ocean (Kikaki et al., 2020; Van Fan et al., 2022). Also, plastic trash poses a hazard to wildlife, as animals can become entangled in it, ingest it, or choke on it, often leading to injury, loss of limbs, or death (Anunobi, 2022). Most of the oceanic plastic trash originates from land, transported either by rivers or eroded from coastal areas, while a smaller portion (~25% of total terrestrial discharge) comes from direct dumping during shipping and fishing activities. Most of these sources are ultimately linked to improperly managed plastic trash, which has significantly increased along with the surge in plastic production over the past 70 years (Zhang et al., 2023).

The most effective solution is to reduce trash generation on land. However, this requires widespread understanding and commitment, which can be challenging to achieve. As such, it's crucial to improve our management of terrestrial trash. A solution that's Sweden and spread to other European countries and even to India is "plogging," a combination of jogging and picking up litter. Originating from Sweden in 2016, the term "plogging" is a blend of the Swedish words "plocka upp" (to pick up) and "jogging." (Raghavan et al., 2022). Several tools have been created to encourage plogging. A noteworthy example is "Data Plogging," a service developed by the non-profit organization ITA Seoul (http://itaseoul.org/). With Data Plogging, volunteers can keep track of the types and locations of trash they collect and store this information online. However, individuals participating in plogging are subjected to repetitive bending, which can lead to increased lumbar loading (Raghavan et al., 2020). Consequently, if people spend extensive time recording and tracking their activities, they may feel exhausted afterward (Martínez-Mirambell et al., 2023). This may make people hesitant to participate in plogging.

To address this issue, we developed a plogging tracking app designed to simplify the process of recording plogging activities. Users can easily record their trash collection efforts by simply taking a photo of the collected trash, eliminating the need for

* Corresponding Author.
Email Address: jungbini@sunmoon.ac.kr (J. B. Lee)
https://doi.org/10.21833/ijaas.2024.03.005
Corresponding author's ORCID profile:
https://orcid.org/0000-0002-8208-0387

manual entry. We conducted a study using the Convolutional Neural Network (CNN) algorithm to create an object detection model that distinguishes objects and people from photos. We utilized the YOLOv5 (You Only Look Once) algorithm (Wu et al., 2021; Wang et al., 2022) and Vision API to automatically detect, predict, and classify the types of photos as well as their qualitative and quantitative quantities when people are plogging. Furthermore, we developed a cross-platform application based on React Native to capture actual trash bags, predict the amount of trash using the deep-learning model, and record it to verify its effectiveness. The application offers enhanced functionality to users by providing basic course tracking, maps, and information on public facilities, thereby presenting users with designated locations for disposing of trash. This application represents a novel contribution to the domain of plogging activity tracking. Processing images of collected trash eliminates the need for manual input, providing a unique differentiator not found in existing manual and labor-intensive methods and services.

In Section 2, we discuss the data collection and training methods for developing the image recognition model. We explain the process of gathering relevant data and preparing it for training the model. We focus on evaluating the performance of the image recognition model in Section 3. In Section 4, we delve into the development and implementation of the mobile app. We provide an overview of the app development process. Finally, in Section 5, we conclude our study by summarizing the key findings and discussing the implications of the developed image recognition model and mobile app in the context of trash management and plogging.

2. Deep learning-based trash recognition model for tracking plogging activity

This study aims to develop a model based on deep learning for trash image recognition and to implement a mobile app that facilitates real-time tracking and recording of plogging activities, thereby enhancing the simplicity and accessibility of monitoring plogging efforts. The development process of this deep learning-based trash recognition model adheres to the procedure outlined in Fig. 1.

2.1. Data collection and labeling

To build a trash recognition model based on a deep learning algorithm, we gathered 3,000 photos of bags filled with trash by conducting web scraping on popular social media platforms, including Instagram, as well as on Google Images. Subsequently, we categorized the collected photos into three groups: 1) photos primarily focused on hands, 2) photos primarily focused on people, and 3) photos primarily focused on trash bags. Fig. 2 plots the labelling results and the distribution of the collected data.

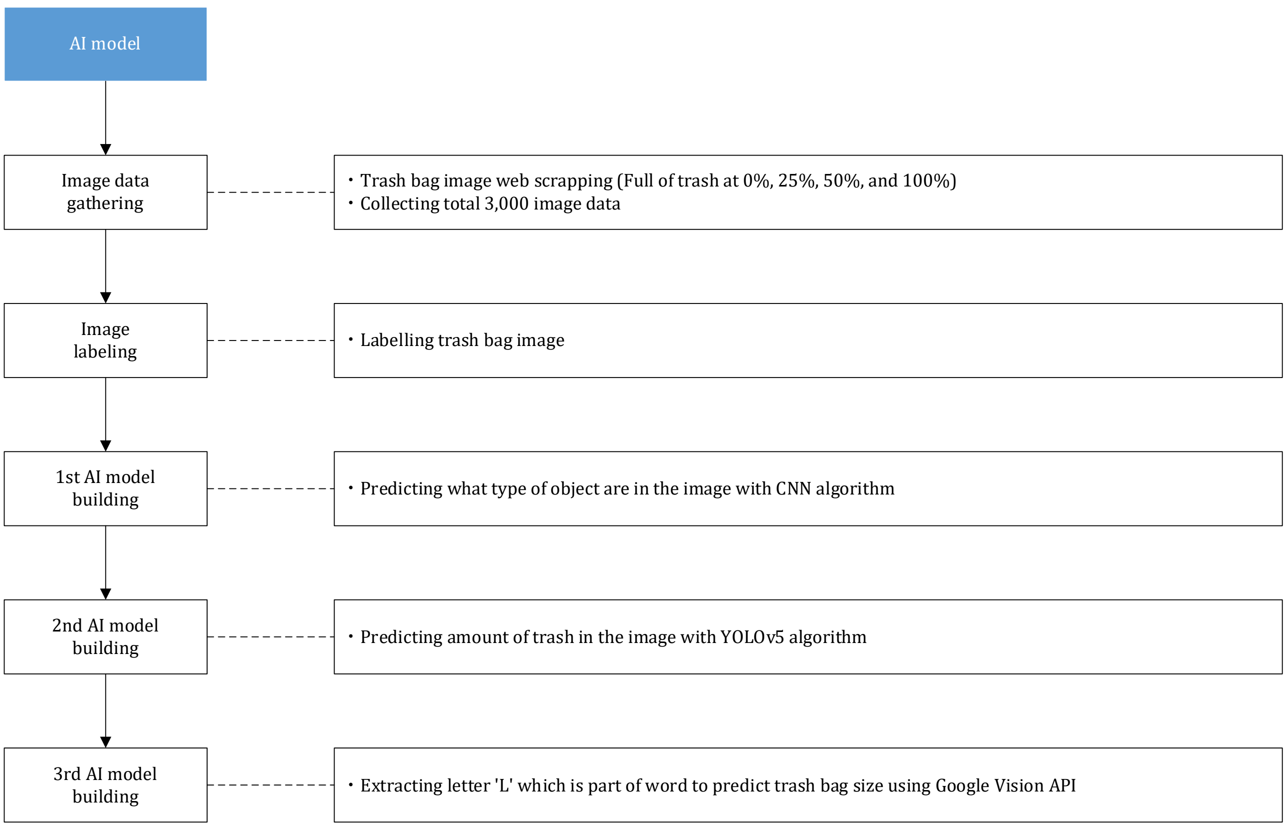

Fig. 1: Process to develop deep learning-based trash recognition model

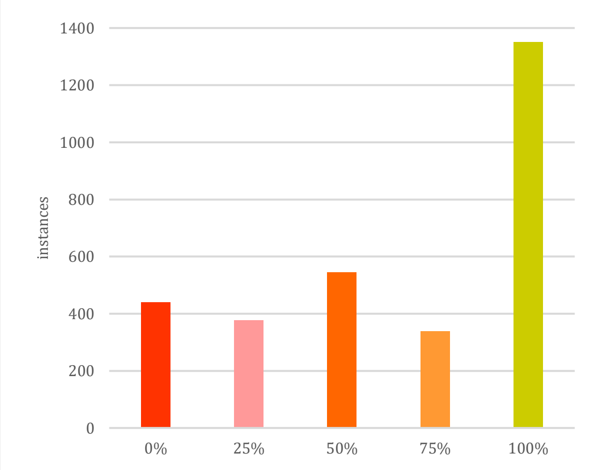

Fig. 2: Distribution of collected images classified into five categories

As shown in Fig. 2, The distribution of the images based on the percentage of trash present shows that images with 100% filled trash bags were more prevalent compared to other percentage categories. These collected images were utilized as training data for both the CNN algorithm and the YOLOv5 algorithm in the subsequent steps of building the AI model.

2.2. Building object classification model from trash images (1st AI model)

The second step involves training a model to identify and classify the main objects in the images using CNN. The objective of this step is to classify and distinguish the primary focus within the image. For example, it determines whether the photo predominantly features hands, people, or trash bags, as mentioned in steps of Section 2.1.

In the preprocessing stage, we loaded 3,000 collected photos and resized them to a specific width and height to be used as input values for the CNN. These photos were then converted into NumPy arrays and stored in a Python list. The labels for the photos to be classified were generated as one-hot encoded arrays. We used the Scikit-learn package in Python to split the training set and validation set in a 7:3 ratio and then saved the split data and labels in Numpy array format using the np.save function. These stored data arrays were then loaded as input values for the CNN model, where we defined the number of categories and classes and normalized the image pixel values to a range between 0 and 1.

Before training the CNN model, we defined it using the sequential API. The Conv2D layer performs 2D convolution operations using 32 filters, each with a size of (3, 3). The size of the input image was also set to have the same output size. We used the ReLU activation function. The MaxPooling2D layer performs maximum pooling operations to reduce the size of a specific map using a pooling window of size (2, 2). In the last Dropout layer, 25% of neurons were randomly deactivated to prevent overfitting, and this layer was repeated twice. The final layer used 128 neurons with the ReLU activation function. Similarly, to prevent overfitting, 75% of neurons were randomly deactivated, and the Dense layer outputted the classification results. We used the Adam optimizer and categorical cross-entropy loss function. Model checkpoints and early stopping callbacks were set up to facilitate the training process. The trained model was evaluated using a validation dataset, and its accuracy was measured.

2.3. Building filled ratio of trash recognition model (2nd AI model)

The second model utilizes the YOLOv5 algorithm to detect trash bags in photos and classify them based on the amount of trash. The reason we used YOLO for image classification is because it's faster than CNN and is specialized for object recognition, specifically targeting the classification of objects like trash bags. We chose YOLOv5 for image classification because it is a state-of-the-art object detection algorithm that is faster and more accurate than its predecessors, such as YOLOv3. YOLOv5 is specifically optimized for real-time object detection, making it ideal for our application, which requires quick and accurate identification of trash bags in images. This choice was also motivated by the need to improve the accuracy of our system in recognizing trash bags from various angles and distances.

The classes represent different levels of trash amount, categorized as 0%, 25%, 50%, and 100%. A total of 3,000 images were labeled and annotated with these class labels. Labeling was done by specifying the bounding box information for the objects and creating a separate class label file. The object information and bounding box information were saved in the YOLO format, and the YOLOv5 model was configured to suit the custom dataset.

2.4. Building capacity of trash bag identification model (3rd AI model)

The fourth step of building the model involves estimating the capacity of the trash bag. In this study, the simplest approach was utilized, using OCR (Optical Character Recognition) of Google Vision to extract only the numbers written on the trash bag as text for initial filtering. In the preprocessing stage, the image size was adjusted, and the image quality was improved through rotation correction and color adjustment. After that, the text area in the image was identified. At this time, a deep learning-based neural network was used. The neural network was trained to recognize main features such as pixel brightness, color, and edges of the image, which then facilitated the detection of text areas.

At this stage, once the text area is detected, it is divided into individual characters, creating separate character images. Based on the spaces between characters, each character is then individually recognized. Subsequently, the segmented character images are converted into text using CNN and RNN. The converted text considers the order and layout of characters, providing accurate text from which numbers are extracted to estimate the liter value from the image. In the assumption of the trash bag's liter value being three digits or more, the first three characters are extracted. If the liter value is less than three digits, the first character is deleted to derive the final liter value as a result.

The capacity measurement of the trash bag was performed specifically for "pay-per-bag" trash bags. However, text extraction varied depending on the image quality and condition. Therefore, an additional step was implemented by extracting the index of the number typically used to indicate size for secondary filtering. Currently, due to a low recognition rate, there is an issue where non-numeric characters are also recognized as numbers, resulting in errors in the output. As a result, the third step results were given less weight in the overall evaluation.

3. Evaluation results

Fig. 3 represents the accuracy of objective classification in trash images (1st AI Model). The accuracy of the training data is approximately 88%, while the accuracy of the validation data is approximately 80%.

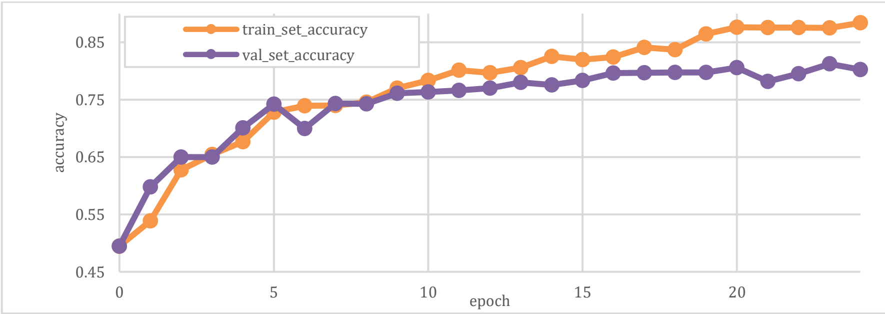

Fig. 3: Performance of object classification mode

Fig. 4 shows the precision, recall, and mean Average Precision (mAP) of Model 2 at different thresholds (0.5 and 0.5 to 0.95) for accuracy as the number of epochs increases during training. The epoch size was set to 100. As shown in Fig. 4, as the epoch size increased, the loss decreased, and the accuracy improved.

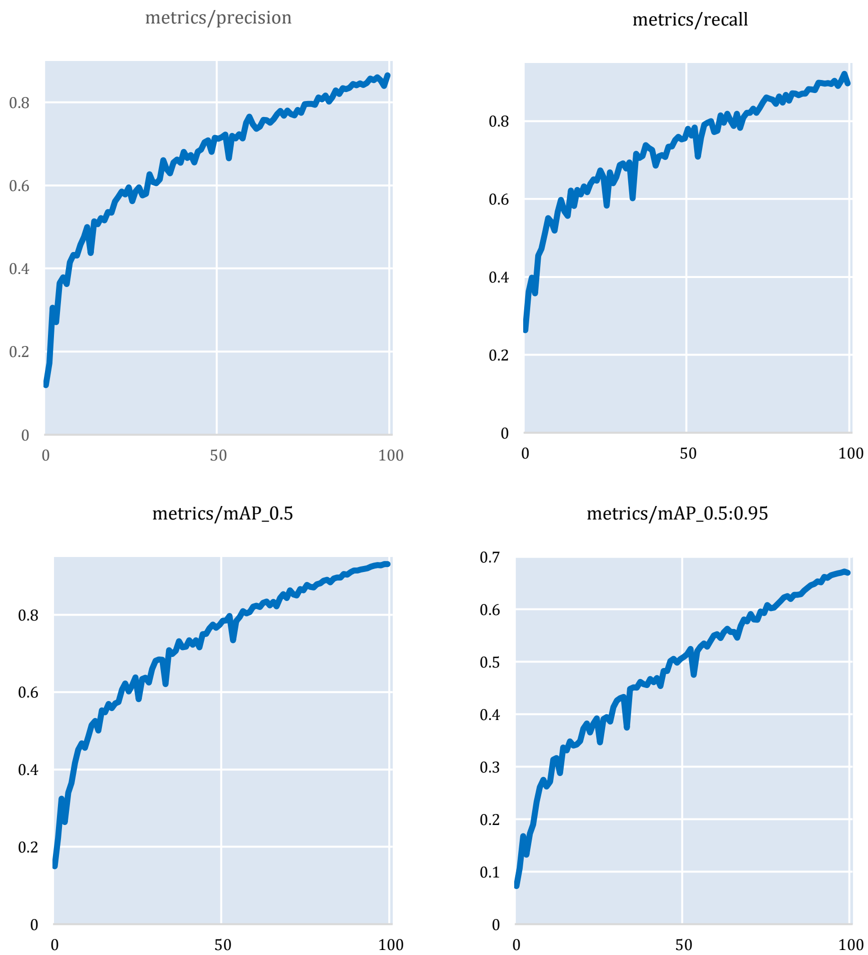

Fig. 4: Performance of filled ratio of the trash recognition model

Fig. 5 depicts the loss rates of the 2nd AI Model. The loss rate of training performance can be explained by dividing it into objectness loss, classification loss, and bounding box regression loss. Objectness loss represents the error in correctly predicting the bounding box at the location of the object. Classification loss represents the error in accurately predicting the class of the object, and bounding box regression loss represents the difference between the predicted bounding box and the actual bounding box of the object. Similar to Fig. 4, the losses decrease as the epoch size increases.

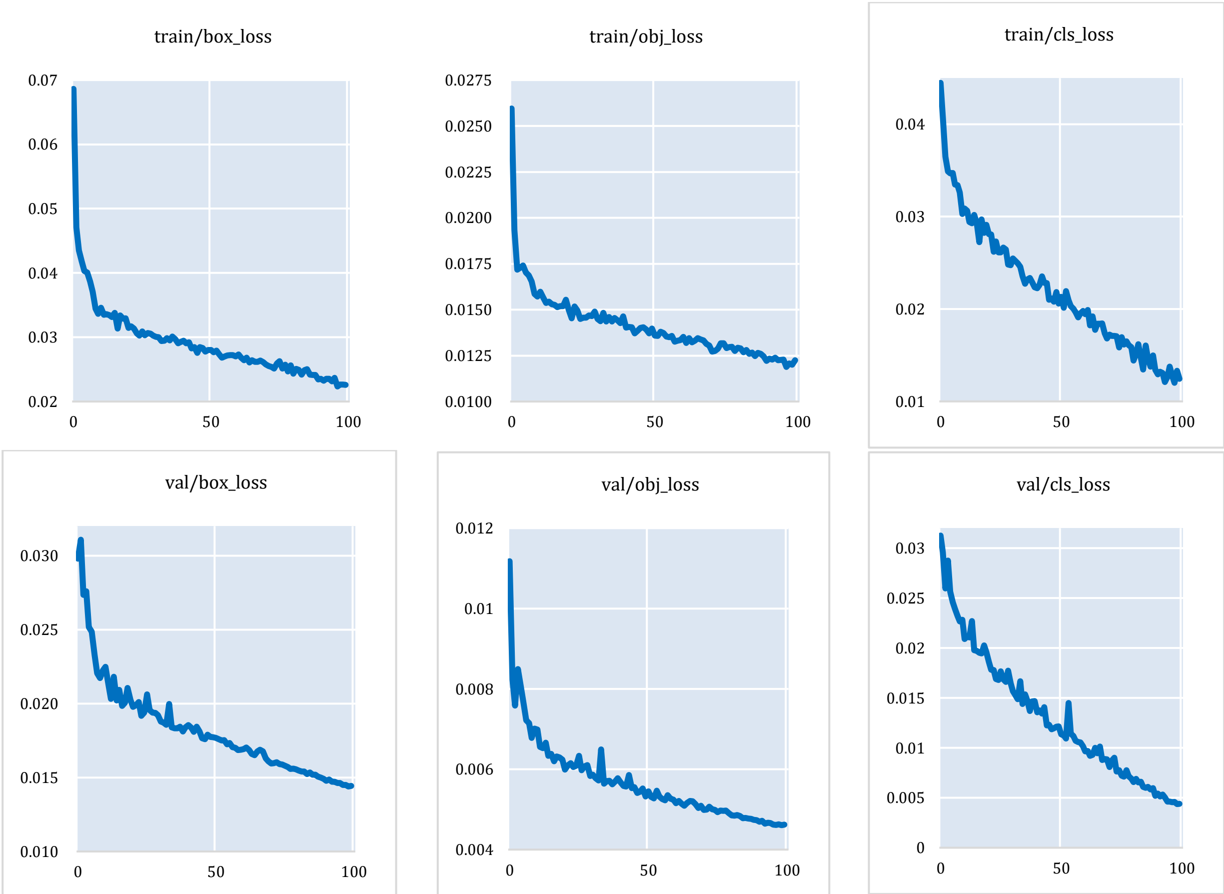

Fig. 5: Loss rates in three categories of the filled ratio of the trash recognition model

Fig. 6 represents the Precision-Recall (PR) Curve for the 2nd AI Model. The PR curve shows the performance of the model at different classes (0%, 25%, 50%, 75%, and 100%), with 97.5%, 90.6%, 90.8%, 88.8%, and 97.9%, respectively. The results indicate that all classes show a higher performance level at a mAP of 0.5, averaging about 93%.

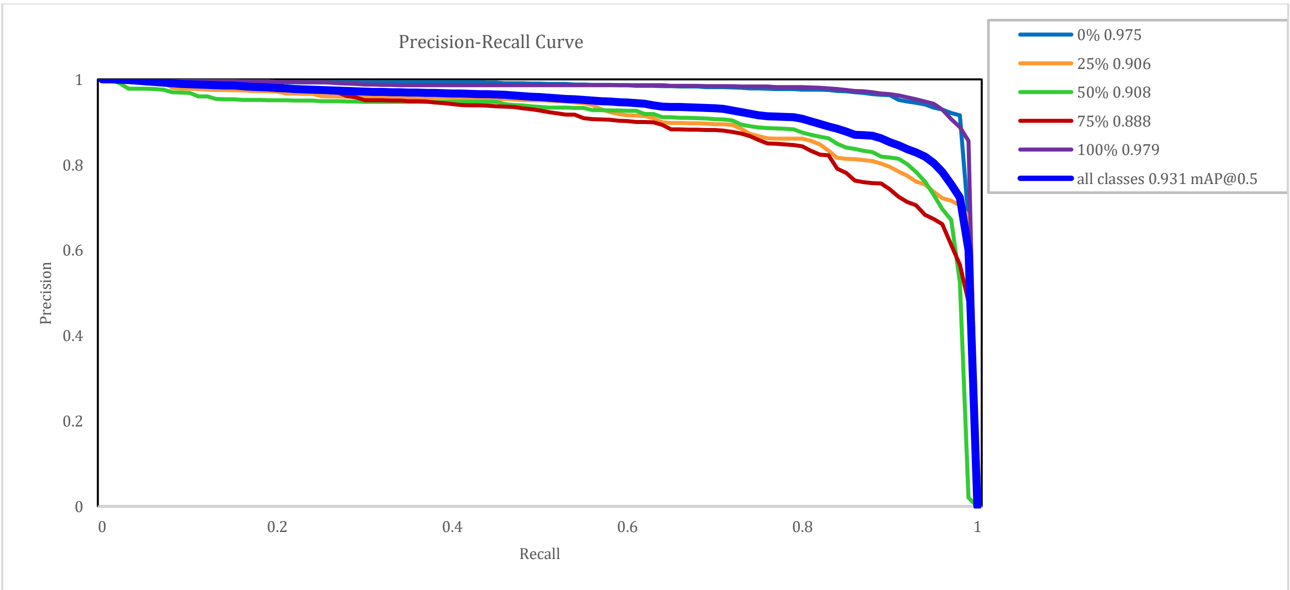

Fig. 6: Precision-recall curve of filled ratio of the trash recognition model

Fig. 7 is an example of a batch image for the final model (3rd AI Model), where an image of a trash bag with visible trash is input into the model, excluding the black trash bag. The model recognizes the bag and displays the percentage of trash filled and the accuracy at the top of the image.

The results of the objects in the image are saved as a text file on the local computer, and the data from this file is then loaded and transmitted to the application server implemented in Section 4.

4. Plogging application implementation

This section introduces the implementation of the plogging application using the deep learning models constructed in Section 3.

4.1. Development environment

Before developing a plogging application, we opted for React Native as a cross-platform framework to ensure compatibility with both iOS and Android operating systems. For the database, we chose MariaDB as an RDBMS (Relational Database Management System). To set up a local server, we utilized Node.js and leveraged Restful APIs. Fig. 8 illustrates the menu structure and flowchart for the main features.

In the React Native-based app, we built a server using Node.js and Express framework, establishing communication with the database through Restful APIs to implement CRUD operations. Additionally, we utilized the Flask framework to create a Python API server, which enabled the execution of deep learning-based models. Fig. 9 illustrates the final system flowchart of the application service.

4.2. User interface

Fig. 10 represents a home screen, the screen for capturing photos after tracking completion and displaying the analyzed results and tracking data. The screen is designed to allow users to operate the app with just one finger, as they may be holding their phone. Service initiation and reset buttons are located in the bottom right corner of the home screen.

Fig. 7: Recognized trash bag images

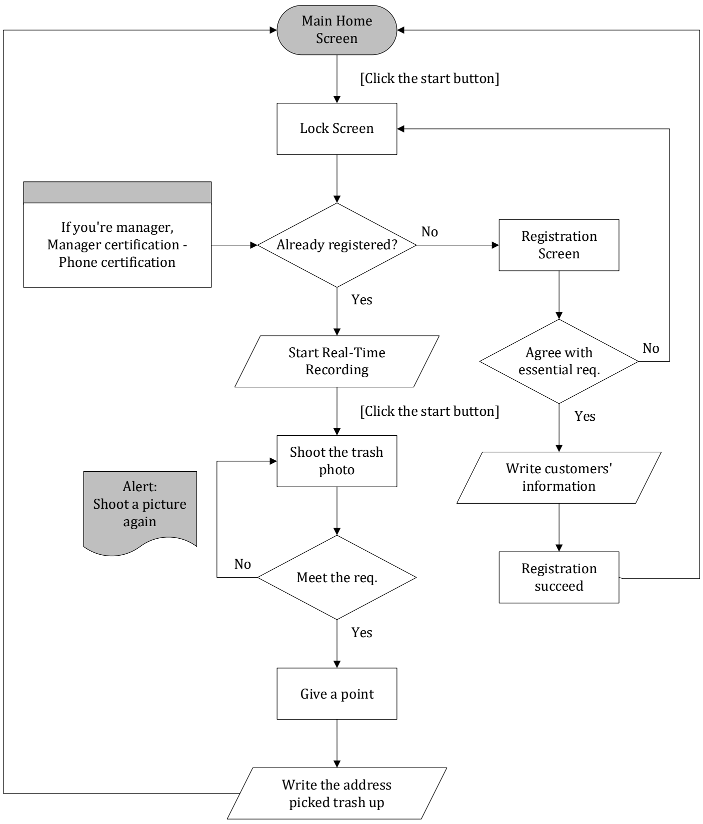

Fig. 8: Flowchart of plogging application

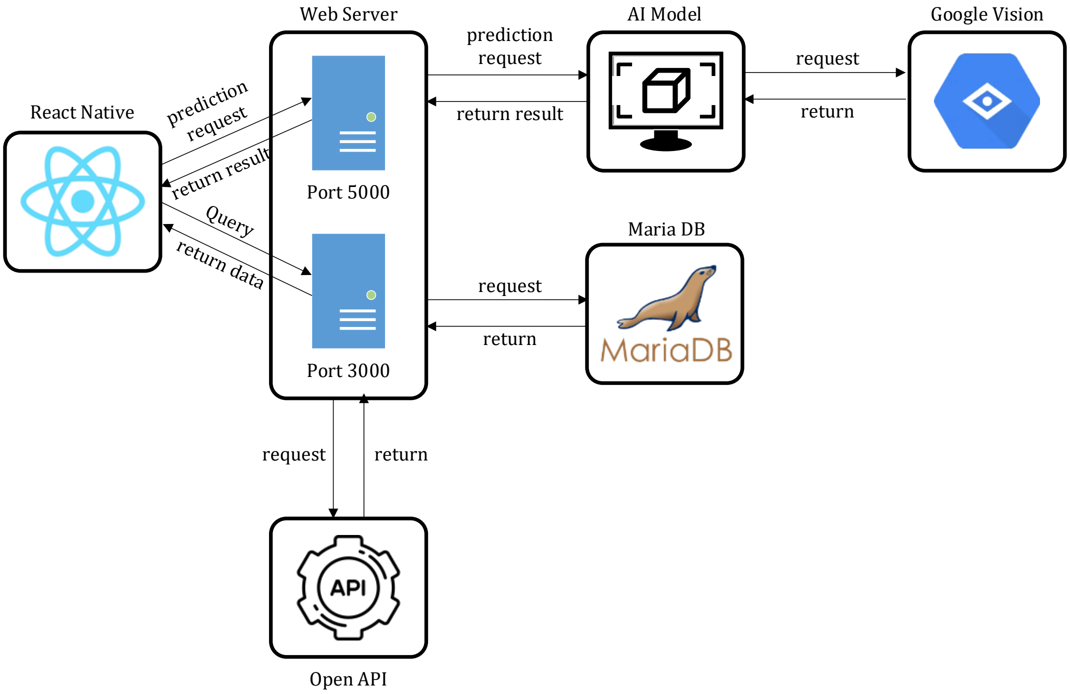

Fig. 9: Process of application services

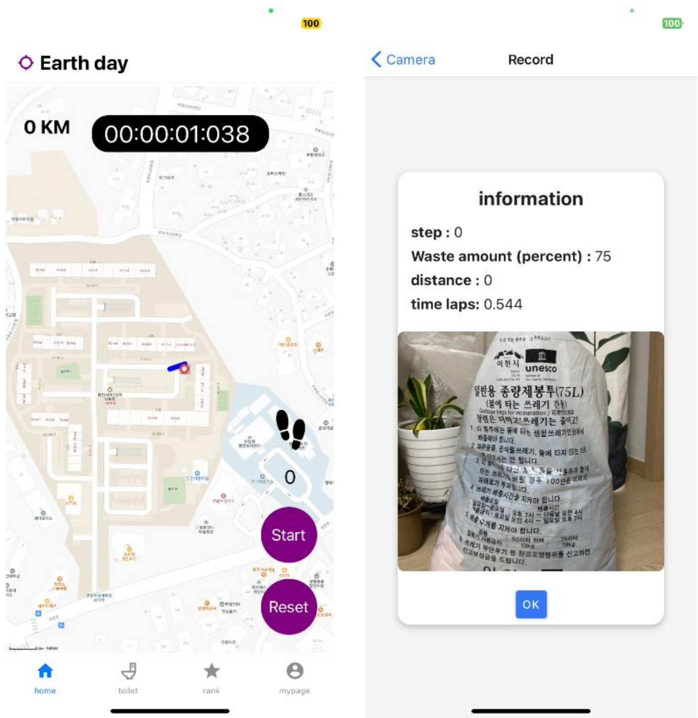

Fig. 10: Home screen UI of application

The home screen displays real-time location, recording time, and step count. To measure distance, we utilized a library that implements the Haversine formula (Fig. 10), which calculates the distance between two sets of latitude and longitude coordinates on the earth's surface (Albahri et al., 2022; Dauni et al., 2019; Diyasa, 2022).

4.3. Application results

As shown in Fig. 10, users can initiate tracking and capture photos when ending the tracking. The captured trash bag images can be analyzed using the deep learning model built in Section 2. If the bag is not accurately recognized, users can take the photo again. After capturing the photo, there is an analysis time of approximately 10-15 seconds, during which an image analysis loading screen is displayed, as shown in the left image of Fig. 11. The image analysis loading screen appears while the image recognition request is sent through the web server, and the deep learning model processes it until the results are received. Once the results are obtained, a result screen is displayed, as shown in the right image of Fig. 11.

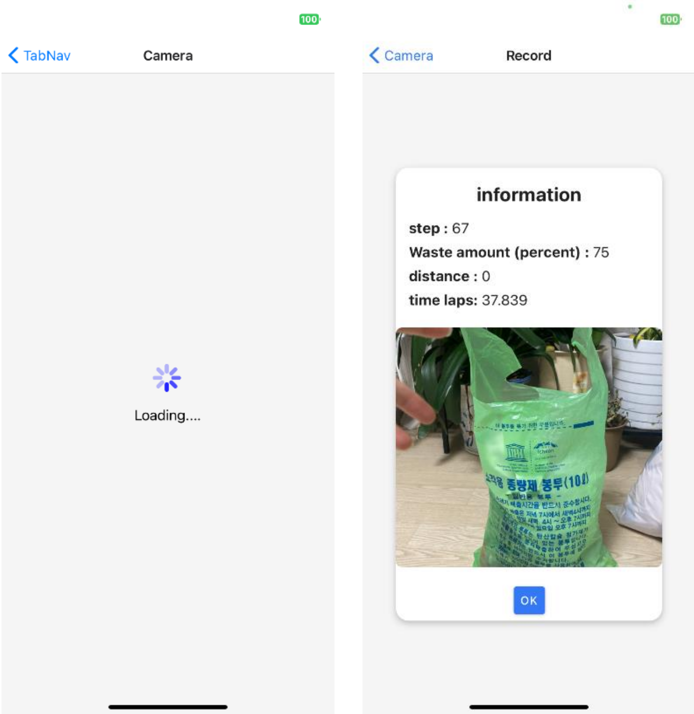

Fig. 11: Image analysis and result screen

When the photo is being analyzed, the system goes through the following process: After passing through the 1st AI Model, a value is generated as "The trash bag is well detected!" Once the photo goes through the 2nd AI Model, the results are stored as image and text values. Fig. 12 provides a result of the image and text saved when passing through the 2nd AI Model.

Fig. 12: Analysis result of the filled ratio of the trash recognition model

The predicted classes (0%, 25%, 50%, 75%, and 100%) are stored among the test results, and the filtered values through the trash bag identification model (3rd AI Model) are transmitted to the web server. The recorded trash quantity in the information shown in Fig. 11 corresponds to the results from the 2nd AI Model, and the final results are stored after passing through the 3rd AI Model.

4.4. Qualitative comparison

Table 1 compares the applications widely used worldwide for tracking plogging. The applications provide GPS-based tracking and gamification services, such as a level system. However, most applications require users to manually record their plogging activities, where the trash was collected, and to recognize the trash photos they have collected.

Clean Swell collects trash data from volunteers worldwide who pick up trash, tagging the data on a map. This is especially helpful in developing countermeasures for handling local trash and tourist-generated trash in resorts and tourist destinations. It is also frequently used domestically. On the other hand, TrashOut provides the world's largest trash map, and users can take photos of the trash they find and register it on the map. Additionally, collecting trash allows users to gain experience points and level up, incorporating a game-like element into the process.

Table 1: Qualitative comparison between plogging applications

Features	Our application	Clean Swell*	TrashOut**
Available Platforms	iOS, Android	iOS, Android	iOS, Android
Location Tracking	Yes	Yes	Yes
Activity Record	Automatic	Manual	Manual
Image Recognition	Automatic	Manual	Manual
Reward	Point and Level	Badge	Point and Level
Community	Yes	Yes	Yes

*https://oceanconservancy.org/trash-free-seas/international-coastal-cleanup/cleanswell; **https://www.trashout.ngo

4.5. Ethical considerations

The images collected for deep learning are publicly available photos of trash bags found on the web. They are not stored in our server but rather converted into a deep learning model for our application. In terms of data privacy, the user passwords and emails managed by the application are encrypted and managed using the SHA-256 hash algorithm. Additionally, the user's location and photos recorded during plogging activities are not shared with other social platforms, and GPS data is used to assess whether the activity is being carried out outdoors.

Furthermore, photos submitted by users of trash bags are used solely for the purpose of awarding points within the system. Lastly, since the application does not recognize anything other than trash bags, other photos are not separately recorded on the server, thereby not posing any ethical issues.

5. Conclusion

In this study, we developed and analyzed a deep learning model for tracking plogging activities to prevent trash from land entering the ocean. We utilized various deep learning and image recognition algorithms to create a high-performance image classification and recognition model. Using this model, we were able to automatically detect and classify the types of photos and quantities of trash when people engage in plogging activities. This introduced AI technology into the existing services, enabling efficient measurement of trash quantities. Furthermore, we developed and implemented a deep learning model-based mobile application that allows users to easily and conveniently identify and collect trash, and we analyzed the results of its application. This app is the first in the field of plogging activity tracking to utilize a deep learning algorithm. Classifying images of collected trash eliminates the need for manual records, significantly providing a unique differentiator not found in existing manual activities and services.

Since the application we developed in our study is the first prototype designed to validate the feasibility of a deep-learning-based recognition model for tracking plogging activities, we were unable to conduct comparative research with similar studies or applications. Additionally, there are limitations in accurately predicting the amount of trash using models trained on images that only capture the front view of the trash. Moreover, the model is biased towards images of 100% filled trash bags, so its performance in classifying other categories, such as 0%, 25%, 50%, and 75%, may be suboptimal. Therefore, for future work, we plan to collect balanced images across all categories of trash bags from various angles and train the deep learning model to enhance its performance. We also aim to conduct comparative research with relevant studies. Moreover, there is a need to update the app's interface and features to enhance user experience and motivate users to engage in plogging. Subsequently, we will conduct a usability evaluation of the application.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (RS-2023-00243114), the MISP (Ministry of Science, ICT Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP (Institute for Information and communications Technology Promotion), and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI22C1472).

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Albahri OS, Zaidan AA, Zaidan BB, Albahri AS, Mohsin AH, Mohammed KI, and Alsalem MA (2022). New mHealth hospital selection framework supporting decentralised telemedicine architecture for outpatient cardiovascular disease-based integrated techniques: Haversine-GPS and AHP-VIKOR. Journal of Ambient Intelligence and Humanized Computing, 13(1): 219-239. https://doi.org/10.1007/s12652-021-02897-4

Anunobi TJ (2022). Hazardous effects of plastic wastes on land biodiversity: A review. Zoologist (The), 20(1): 80-86. https://doi.org/10.4314/tzool.v20i1.10

Dauni P, Firdaus MD, Asfariani R, Saputra MIN, Hidayat AA, and Zulfikar WB (2019). Implementation of Vaversine formula for school location tracking. Journal of Physics: Conference Series, 1402(7): 077028. https://doi.org/10.1088/1742-6596/1402/7/077028

Diyasa IGSM, Prasetya DA, Idhom M, Sari AP, and Kassim AM (2022). Implementation of Haversine algorithm and geolocation for travel recommendations on smart applications for backpackers in Bali. In the 2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), IEEE, Jakarta, Indonesia: 504-508. https://doi.org/10.1109/ICIMCIS56303.2022.10017760

Kikaki A, Karantzalos K, Power CA, and Raitsos DE (2020). Remotely sensing the source and transport of marine plastic debris in Bay Islands of Honduras (Caribbean Sea). Remote Sensing, 12(11): 1727. https://doi.org/10.3390/rs12111727

Martínez-Mirambell C, Boned-Gómez S, Urrea-Solano M, and Baena-Morales S (2023). Step by step towards a greener future: The role of plogging in educating tomorrow's citizens. Sustainability, 15(18): 13558. https://doi.org/10.3390/su151813558

Raghavan R, Panicker VV, and Emmatty FJ (2020). Posture based assessment of plogging activity. In the 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), IEEE, Pondicherry, India: 1-5. https://doi.org/10.1109/ICSCAN49426.2020.9262447 **PMid:33096533 PMCid:PMC7309575**

Raghavan R, Panicker VV, and Emmatty FJ (2022). Ergonomic risk and physiological assessment of plogging activity. Work, 72(4): 1337-1348. https://doi.org/10.3233/WOR-205210 **PMid:35723137**

Van Fan Y, Jiang P, Tan RR, Aviso KB, You F, Zhao X, and Klemeš JJ (2022). Forecasting plastic waste generation and interventions for environmental hazard mitigation. Journal of hazardous materials, 424: 127330. https://doi.org/10.1016/j.jhazmat.2021.127330 **PMid:34600379**

Wang Z, Jin L, Wang S, and Xu H (2022). Apple stem/calyx real-time recognition using YOLO-v5 algorithm for fruit automatic loading system. Postharvest Biology and Technology, 185: 111808. https://doi.org/10.1016/j.postharvbio.2021.111808

Wu W, Liu H, Li L, Long Y, Wang X, Wang Z, and Chang Y (2021). Application of local fully convolutional neural network combined with YOLO v5 algorithm in small target detection of remote sensing image. PLOS ONE, 16(10): e0259283. https://doi.org/10.1371/journal.pone.0259283 **PMid:34714878 PMCid:PMC8555847**

Zhang Y, Wu P, Xu R, Wang X, Lei L, Schartup AT, and Zeng EY (2023). Plastic waste discharge to the global ocean constrained by seawater observations. Nature Communications, 14(1): 1372. https://doi.org/10.1038/s41467-023-37108-5 **PMid:36914656 PMCid:PMC10011382**