Appendix 2

Result of Demonstration Test for Beach Litter Survey Using Stationary Cameras

1. Purpose of the demonstration test

The purpose of the demonstration test is to ensure the practicality of Annex Section I and II of the guidelines with the following elements (Table 1).

Table 1. Points to ensure practicality

Concrete-ness	The survey will be conducted by inexperienced staff under expert supervision to organize the items to be included in the guideline.
Versatility	Note that the contents of the guidelines need to be generalized so that they can be used as a reference even when the equipment used or the research environment is different. The image analysis method will utilize AI while also examining the possibility of manual detection.
Novelty	Verify whether the guideline can solve the problems of efficiency, accuracy, and reproducibility of the existing surveys to be verified.

2. Demonstration test outline

Table 2 shows the outline of the demonstration test.

Table 2. Demonstration test outline（1/2）

Items	Outline
Survey period/ frequency	Installation of the stationary camera: 29th June 2024 Monitoring period: 29th June 2024 - 31st December 2024 Camera setting: AM5:00 – PM7:00 every day, photographed once an hour On-site visual inspection: September and collection survey: 9th September 2024
Survey points	Rokudoji beach in Imizu city, Toyama, Japan (Sandy beach)
Survey area	70.9 m^2 (see figure)
Survey method (See Appendix B for detailed information)	The amount and type of beach litter was investigated and analyzed using the following four methods: I Stationary camera survey with automatic detection of beach litter by AI (semantic segmentation) II Stationary camera survey with manual detection of beach litter from images III On-site visual inspection (implemented with reference to the "Index Evaluation method for Waterside Scattered Debris (Coastal Version)" (Tohoku Regional Development Bureau, Ministry of Land, Infrastructure, Transport and Tourism, etc.)) IV Collection survey (with reference to the “Guidelines for Investigating the Composition of Drifted Debris for Local Governments (June 2023, 3rd edition)”) * Methods III and IV can be performed by the general public with the guidance of an experienced person. Furthermore, the same piece of litter in the front and back of the image will appear different in size (pixels) on the image, due to the difference in GSD (ground sample distance). Then, the survey area was divided into four sections (Section 1, 2, 3, and 4) as the reference for discussions. The sections refer to 0-25%, 25-50%, 50-75%, and 75-100% from the front of the camera's field of view, respectively in Method II and IV (Figure 4).

Table 3. Demonstration test outline（2/2）

Items	Outline
Survey target (See Table 4 for detailed information)	Ⅰ Number of pixels and covered area of beach litter Ⅱ Number of pixels, number of pieces, volume of beach litter Ⅲ Volume of beach litter Ⅳ Number of pieces, volume, and weight of beach litter － To estimate the litter covered area, positional information correction was conducted on the taken image by projection transformation using surveying technology. － The target in this survey was only man-made objects, although the amount of wood (lumber, etc.) from man-made objects was excluded because the item is difficult to distinguish from driftwood (natural objects). － In Survey I, only the total amount of man-made objects is obtained since the AI cannot perform detailed waste sorting. － The types of items to be surveyed in Survey II, III, and IV are plastic bottles, food containers (incl. fast food cups, etc.), plastic bags (shopping bags, food bags, industrial bags, etc.), seedling pots, and other man-made litter, based on the characteristics confirmed in the preliminary survey.
	Dr. Kako's lab, Japan NUS Co., Ltd. (JANUS), Northwest Pacific Region Environment Cooperation Center (NPEC)
Image processor/analyst	Dr. Kako's lab, Japan NUS Co., Ltd. (JANUS)
Method of organizing results	－ Time series data on the number of pixels and the covered area of beach litter (Survey I) － Comparison of the numbers of pixels of beach litter estimated (Method I and II) － Comparison of the number of pieces of beach litter estimated (Method II and IV) － Comparison of the volume of beach litter estimated (Method II, III, and IV) － Verification of human error in visual surveys (Method II) － Comparison of time spent on surveys (Method I - IV) － Comparison of survey cost (Method I - IV)

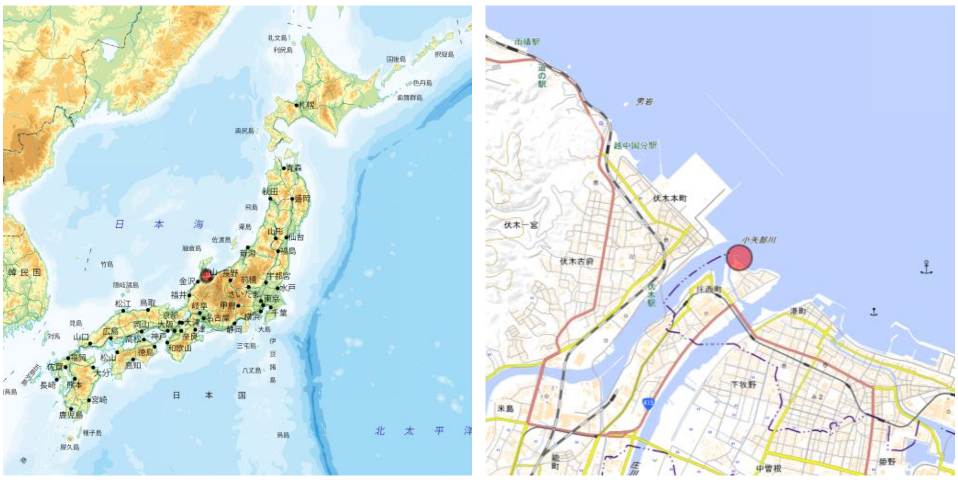


Figure 1. Location of Rokudoji beach in Imizu city, Toyama, Japan

（Source: The Geospatial Information Authority of Japan (GSI) map）

Figure 2. Location of the stationary camera installed

（Source of the left figure: GSI map）

Figure 3. Camera shooting coverage (red frame)

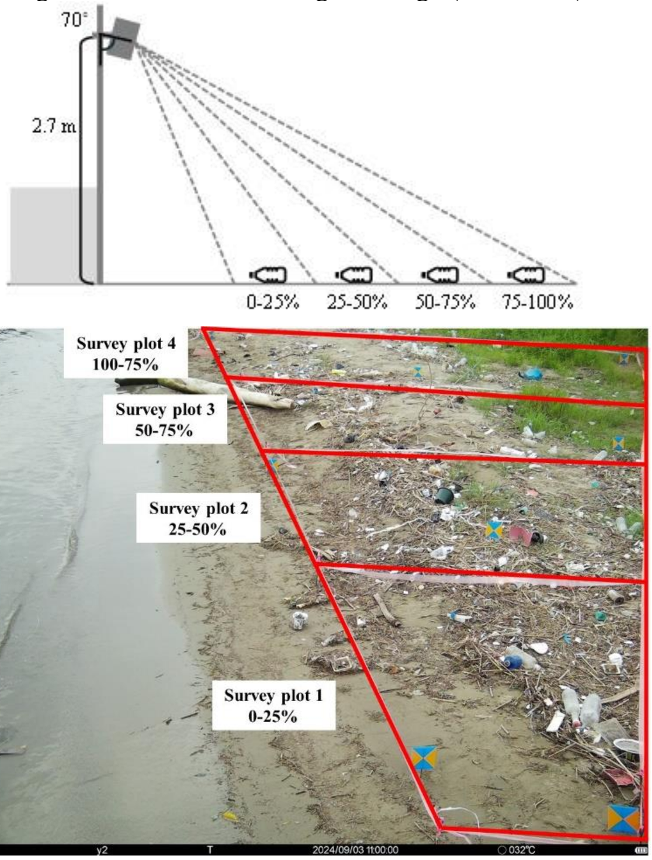


Figure 4. Altitude and angle of camera, and the survey plots (red frame)

Table 4. Survey methods and litter information obtained by each method

Items	Survey Method	Litter information					Purpose
		Pixels	Covered Area	Number of Pieces	Weight	Volume	
Stationary camera survey	I Stationary camera and AI detection						– The number of pixels is calculated by automatically detecting litter from the stationary camera images using AI. – The litter covered area is calculated by projection transformation based on the pixels of the survey results.
	II Stationary camera and manual detection						– The number and types of litter are estimated by visually identifying from the stationary camera images, which are difficult to be automatically detected by AI (Method I). – The estimated number of pixels obtained by Method II is compared to the results of Method I to confirm the accuracy of AI.
Manual survey	III On-site visual inspection						– The estimated volume of beach litter and human error is compared to that of Method II.
	IV Collection survey (ground truth)						– The types and number of litter, volume of litter are compared to the result of Method II.

3. Results of the demonstration test

3.1 Analysis of time series data of beach litter

For Method I (stationary camera and AI detection), time series images for the period from July 1 to December 31, 2024 were obtained, and AI was used to detect litter in images taken at 12:00 every day, and the number of pixels was calculated.

When the beach is photographed from an angle using a stationary camera, the litter in the front appears larger and the litter in the back appears smaller in the image, and there is a difference in the number of pixels. In order to correct this imbalance and more accurately represent the amount of litter, the oblique images were converted into a top-down view by projection transformation based on the results of the surveying (See Figure 5), followed by calculating the litter covered area based on the number of pixels (See

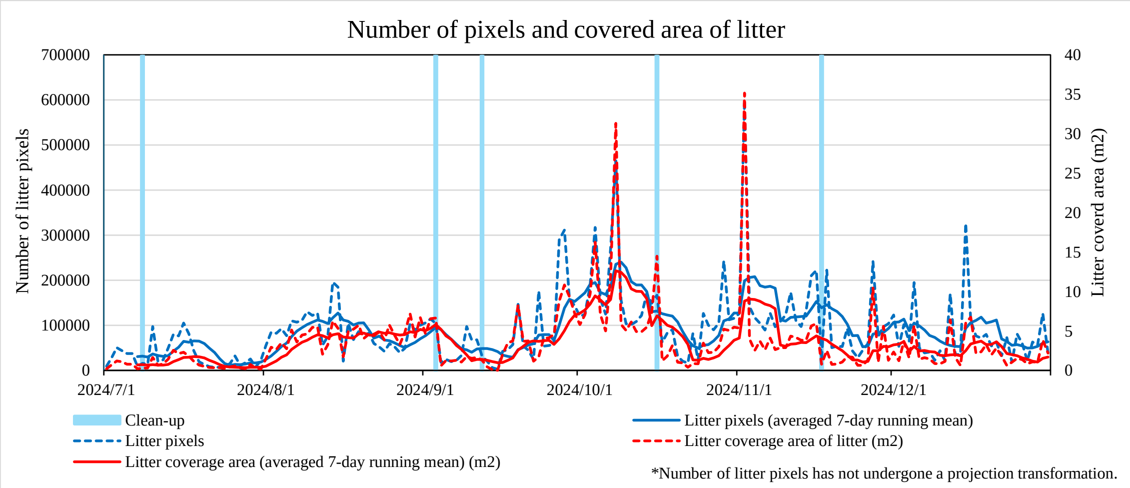
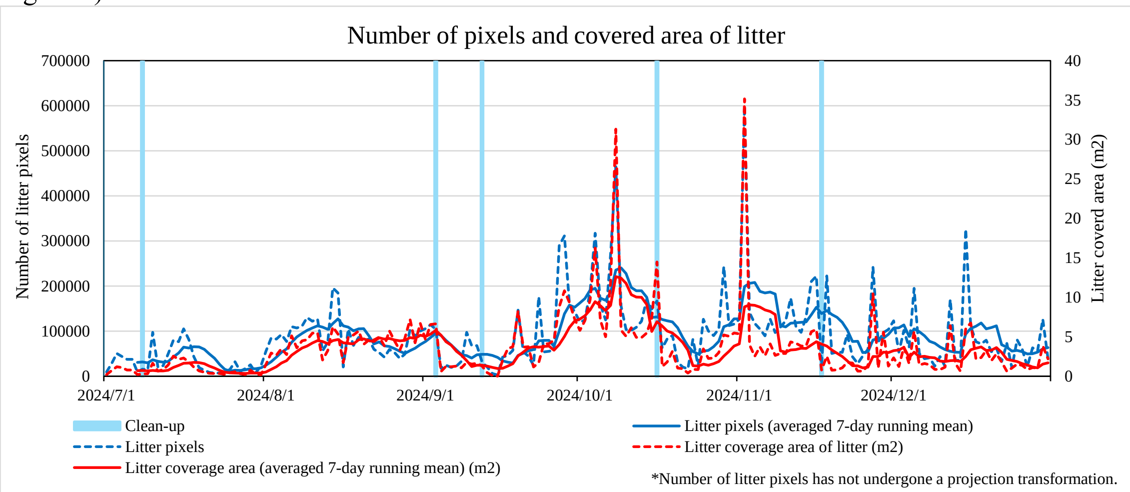


Figure 6).



We included several meteorological data that was assumed to be related to wind, rain, and water levels, as well as the dates on which cleaning activities were conducted, and examined the relationship with the increase or decrease in litter. Although there were several local events being held in the area around the camera installation site and in the upper reaches of the river during the survey period, none of these appeared to be contributing to the large amount of litter in the environment. The results and discussion are shown below.

- The correlation coefficient between the time series variation of the litter covered area and the time series variation of the number of pixels that have not undergone a projection transformation was significant ($p < 0.01$), so it is thought that it is possible to grasp the approximate time series variation in the amount of litter from the number of pixels of litter identified from the number of pixels (that have not undergone projection transformation) (See Figure 7).
- Some data was thought to have abnormally large values of pixels due to the sun glint or fog of the camera lens being mistakenly detected. Therefore, data with more than 300,000 pix was considered an outlier, and 7-day running means were created with and without the outliers. As

the graphs for both sets of data almost overlapped and the overall trend was almost the same, it is thought that taking a 7-day running means can reduce the impact of outliers (See Figure 8). In this demonstration, a 7-day moving average was used, but the appropriate number of days for averaging may differ depending on the characteristics of regions, such as the frequency of rainfall.

- The number of litter pixels is greatly reduced after clean-up activities. When the analysis was limited to the period when there was no impact from the clean-up activities (July 9 to September 2), there was a high correlation between the number of litter pixels and wind speeds of the easterly and northerly wind components. The correlation coefficients are both significant ($p < 0.01$). This indicates that the higher the wind speed of the northward and easterly components, the greater the amount of litter, and this is thought to be due to the shape and direction of the coast and the breakwater blocks (See from Figure 9 to Figure 11).
- It was not possible to confirm a clear relationship between metrological data (duration of high tide, rainfall, preceding dry days*, daily sunshine duration, air pressure, Oyabe River water level) other than wind speed and the number of litter pixels (see from Figure 20 to Figure 25). It is possible that the rainfall upstream of the river has a greater impact than the rainfall at the camera installation site. There may also be a lag between the outflow and the washing up of litter.

* Preceding Dry Days: Refers to the number of consecutive days during which the daily precipitation falls below a certain threshold (e.g., less than 10mm per day). The threshold may vary depending on the region.

Figure 5. Image of the beach taken by a stationary camera (left) and image after projection transformation (right)

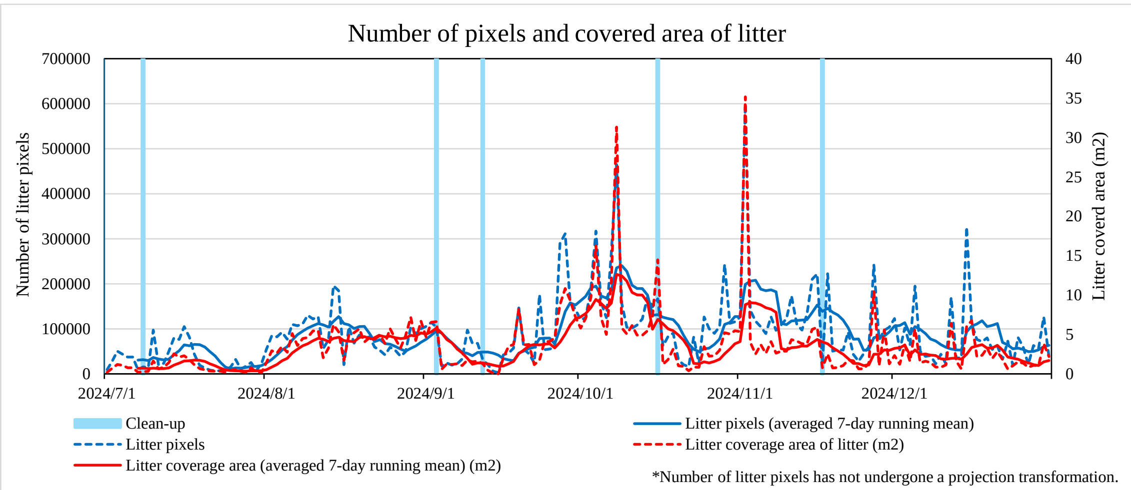


Figure 6. Time series data: Number of pixels and litter covered area

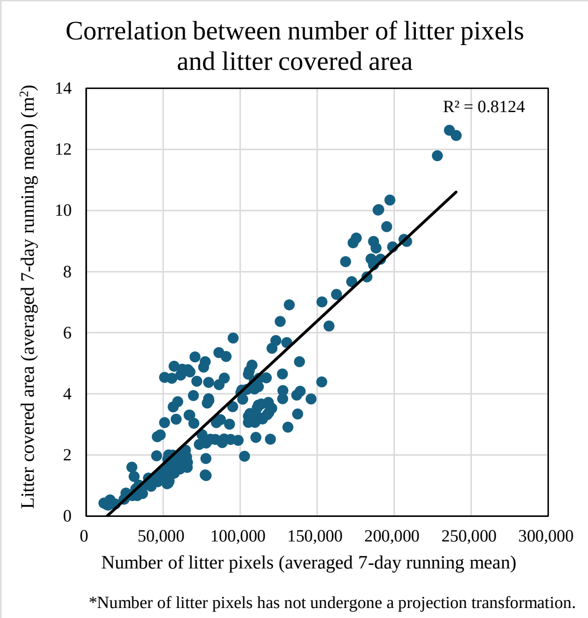


Figure 7. Correlation between number of litter pixels and litter covered area

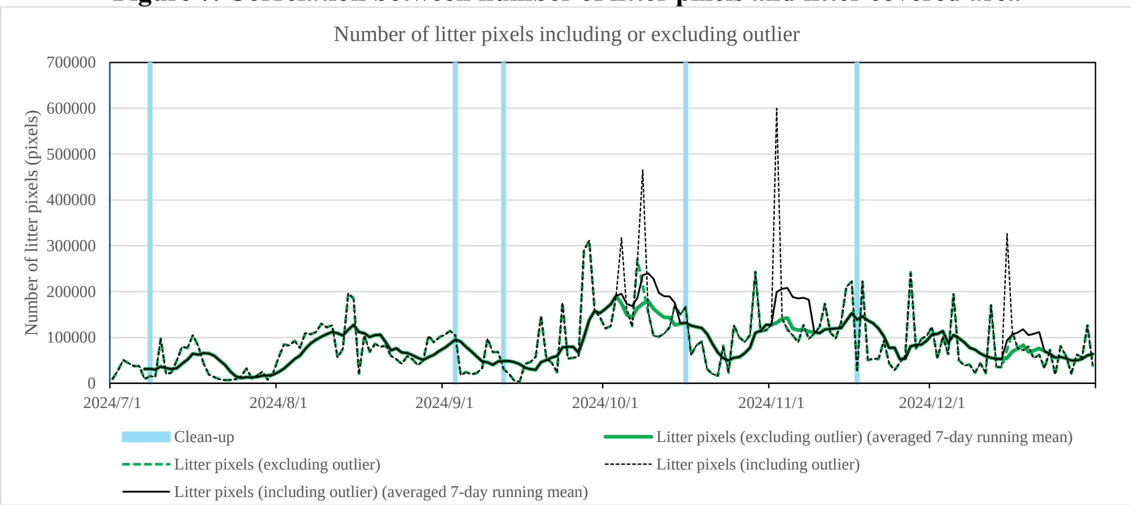


Figure 8. Time series data: Number of litter pixels including or excluding outlier

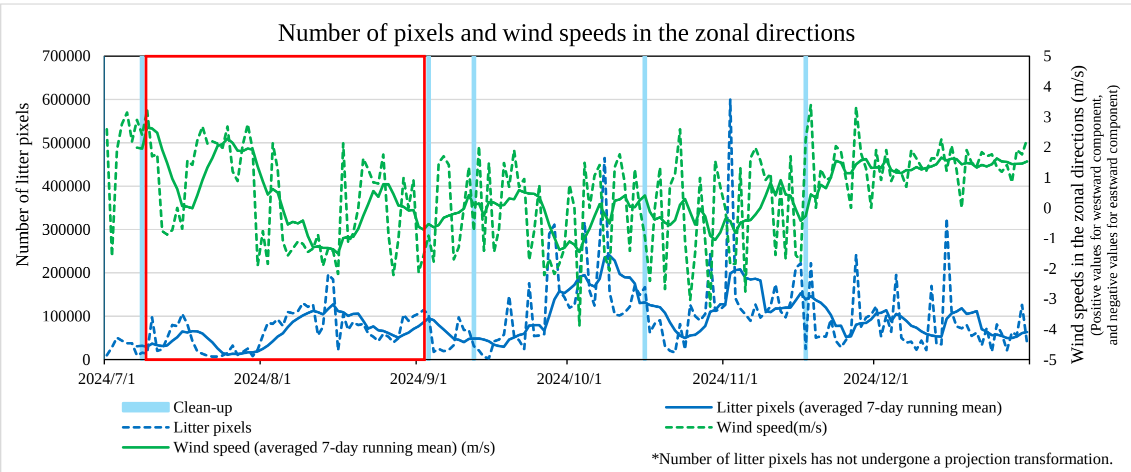


Figure 9. Time series data: Number of pixels and wind speeds in the zonal directions

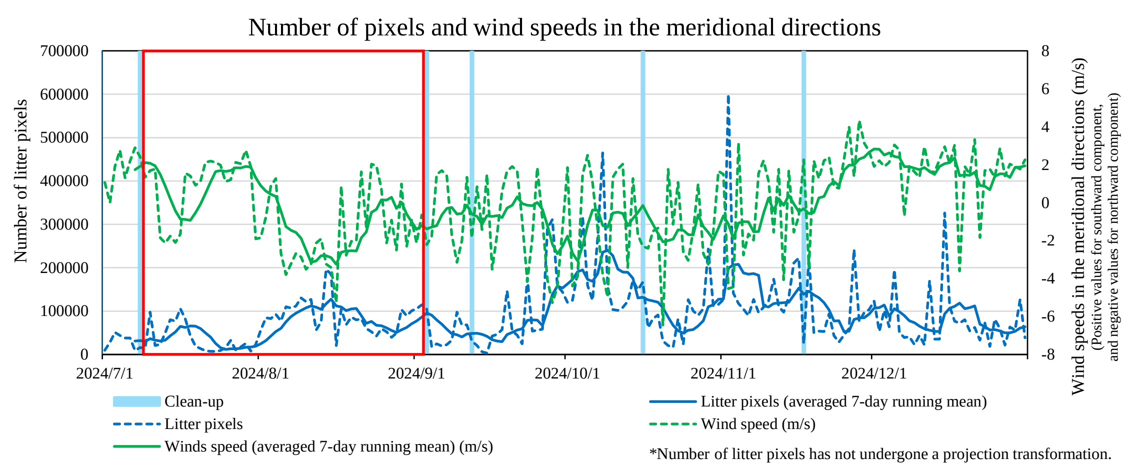
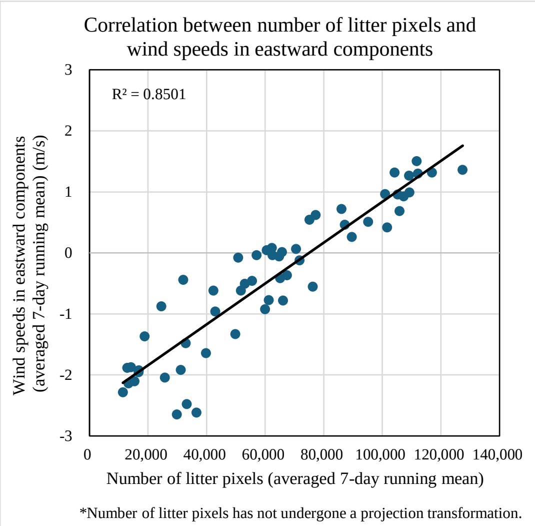


Figure 10. Time series data: Number of pixels and winds speeds in the meridional directions



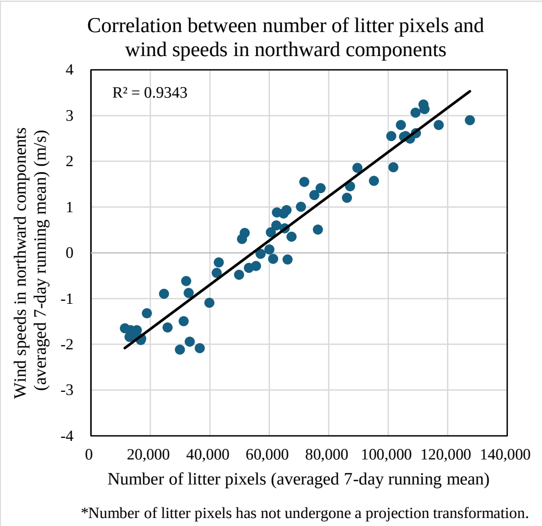


Figure 11. Correlation between number of litter pixels and wind speed: eastward components (left) and northward components (right)

3.2 Results of comparisons of the amount of beach litter estimated by each survey method

On September 3, 2024, Method I (stationary camera and AI detection), Method II (stationary camera and manual detection), Method III (on-site visual inspection), and Method IV (collection survey (Ground truth)) were conducted simultaneously and the results were compared. For Methods I and II, we used images of days with a lot of litter (clean-up activities) and days with little litter (after clean-up activities) on days other than September 3, and compared the number of litter pixels between the methods using a total of 6 images.

3.2.1 Comparison of the survey results in number of pixels and covered area of litter (Method I and II)

We compared the estimated number of litter pixels of 6 images using Method I (stationary camera and AI detection) and Method II (stationary camera and manual detection). Examples of the detection result images of Method I and II are shown in Figure 13. and Figure 14.
As a result, the number of pixels detected in Method I and those of Method II showed a similar trend, suggesting that detection by AI produces results similar to those of the human eye, and that AI can be used as a substitute for human detection.
On September 14, the number of pixels of Method I was relatively high, whereas that of Method II was very low. This is considered to be because the AI mistakenly detected sun glint on the sea surface as litter (See Figure 15).
In addition, the AI detected a larger number of pixels as litter overall. One possible reason for this is that the AI detected as litter what researcher α and β did not detect as litter, such as very small pieces of litter or litter in the back of images with low resolution.

Table 5. Results of number of pixels and litter covered area obtained by Method I (stationary camera and AI) and Method II (stationary camera and manual detection)

Number of Pixels	Stationary Camera and AI (pix)	Stationary Camera and Manual Detection (pix)		
		Researcher α	Researcher β	Average
Jul. 12 (After clean-up)	159,474	52,538	75,799	64,169
Sep. 3 (Before clean-up)	576,002	509,706	415,502	462,604
Sep. 14 (After clean-up)	773,943	22,613	17,295	19,954
Oct. 15 (Before clean-up)	848,329	600,136	718,120	659,128
Oct. 17 (After clean-up)	464,008	164,757	194,764	179,761
Nov. 16 (Before clean-up)	1,010,172	716,073	926,540	821,307

Note:
These results include the uncertainty associated with each estimate, including detection of non-litter as litter or failure to detect litter.

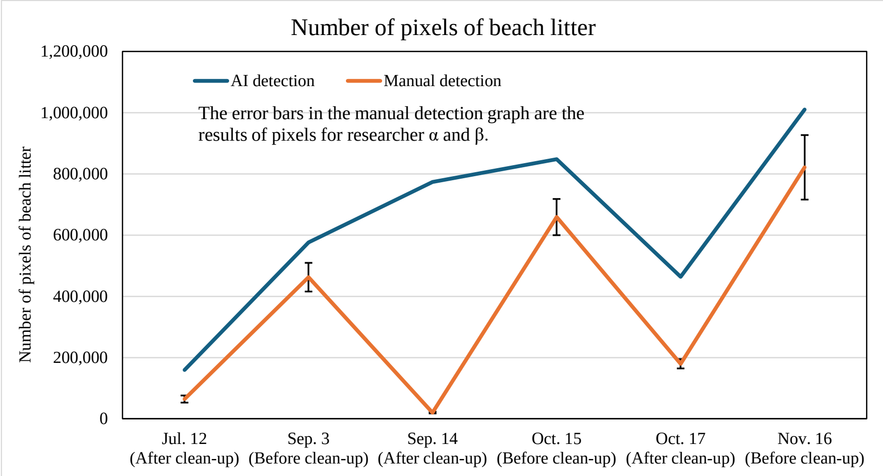


Figure 12. Comparison of the survey results in number of litter pixels of Method I and II

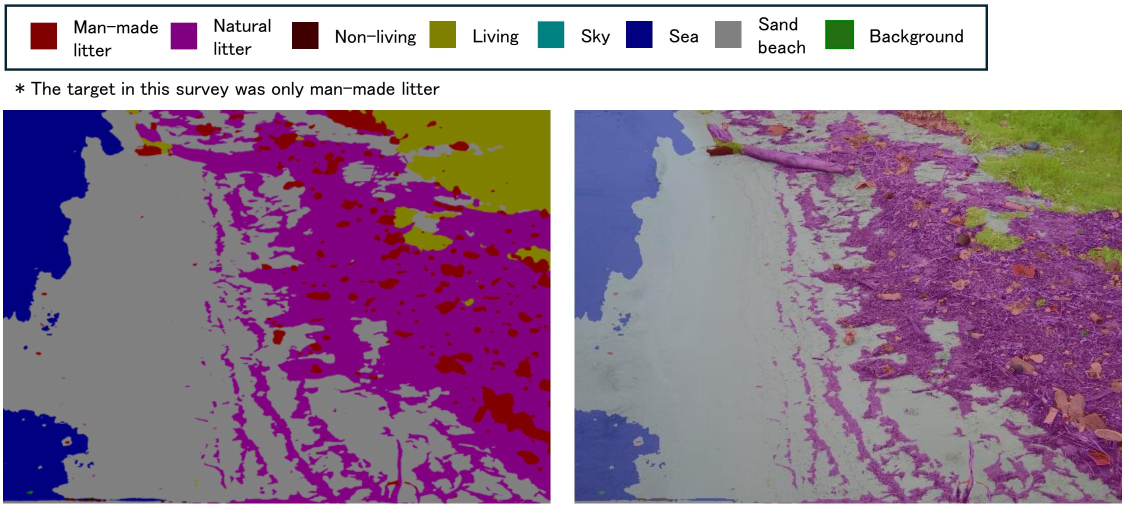
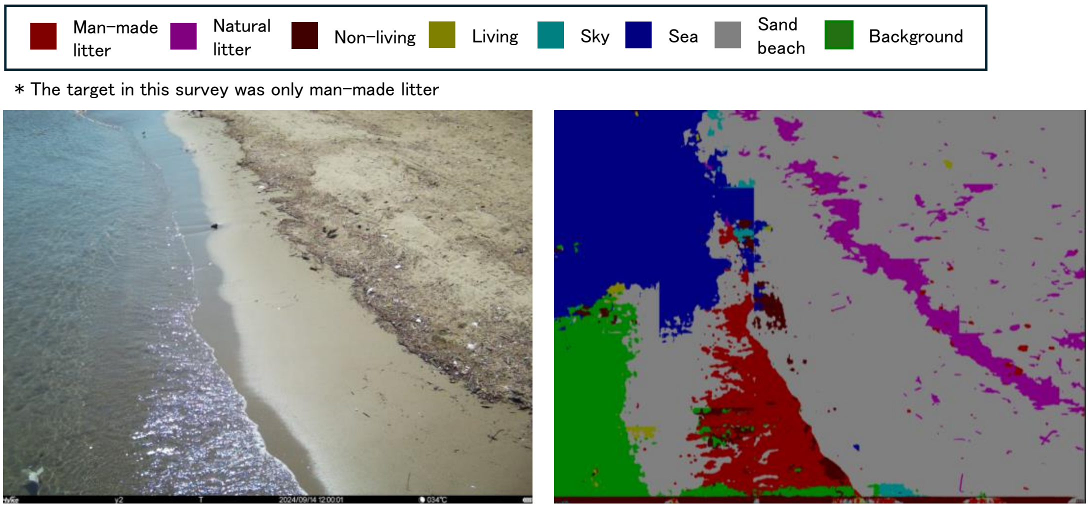


Figure 13. Results of the AI analysis of an image (September 3rd, 2024): Result of the AI detection (left) and Result of the AI detection superimposed on the original photograph (right)

(Application used: Labelme v5.5.0 (the Massachusetts Institute of Technology (MIT))

Figure 14. Result of the manual detection from an image (September 3) : Result of the detection by researcher α (left) and Result of the detection by researcher β (right)



(Application used: Labelme v5.5.0 (the Massachusetts Institute of Technology (MIT))

Figure 15. False positives in image detection using AI (September 14): Original image (left) and Result of the AI detection (right)

3.2.2 Comparison of the survey results in pieces (Method II)

The number of litter items using Method II (stationary camera and manual detection) was compared with the ground truth obtained using Method IV (beach litter collection) (see Table 6), and the detailed results of the survey, which was divided into four sections from the front to the back of the image, are shown in Table 15 to Table 18.

As a result, the number of items detected was underestimated compared to the ground truth. The litter detection rates for plastic bottles and other man-made waste were around 50%, while the rates for food containers, plastic bags, and seedling pots were around 30% or less. The possible reasons for this are as follows:

- The litter in the front of the image is easier to identify, while the litter in the back is more difficult to identify, so the detection rate decreases, leading to an underestimation (see Table 15 to Table 18).
- A stationary camera can only detect objects that are visible on the surface, but the ground truth also includes litter below obstacles.
- Detection of the food containers, plastic bags, and seedling pots was considered to be difficult because their average volumes were 0.15 m^3/piece, 0.4 m^3/piece, and 0.22 m^3/piece, respectively, which is smaller than the average volume of plastic bottles (0.56 m^3/piece), and they are also irregularly shaped (see Figure 16 and Figure 17). On the other hand, the detection rate for other man-made objects was the second highest (0.46) after plastic bottles, even though their average volume was small (0.14 m^3/piece). It is considered that many pieces of litter are classified as "Other man-made objects" because, although it is relatively easy to detect that the objects in the background are artificial, the images are unclear, making further classification impossible.

Table 6. Results of piece count comparison between stationary camera and manual detection (Method II) and the ground truth (unit: pieces)

Objects	Stationary Camera and Manual Detection (Number)			Ground Truth (Number)	Detection Rate (Average of Manual Detection / Ground Truth)
	Researcher α	Researcher β	Average		
Plastic bottles	21	18	20	35	0.56
Food containers (incl. fast food cups, etc.)	9	6	8	74	0.10
Plastic bags (for shopping, food, business use, etc.)	0	0	0	1	0
Seedling pots	6	3	5	14	0.32
Other man-made objects	244	209	227	490	0.46
Total	280	236	258	614	0.42

Note:
1. Estimations of the stationary camera and manual detection are the average of the survey results from each of the two researchers.
2. See II-III in Appendix B for the methods to conduct surveys and estimates.

Figure 16. Plastic bottles (left) and food containers (right)

Figure 17. Seedling pots (left) and other man-made litter (right)

3.3 Comparison of time spent on survey

Table 7. Time spent on survey (Time required for the first survey)

Survey Methods		Data Sampling (Unit: Hour)		Data Processing/Analysis (Unit: Hour)		Total Person-Hour (Unit: Person*Hour)
		Preparation	Measurement /Sampling	Processing	Detection Classification Quantification	
I	Stationary camera and AI	4 (2 workers)	0	0	0	8
II	Stationary camera and manual detection				0.75 (1 worker)	8.75
III	On-site visual inspection	1.5 (2 workers)	0.5 (2 workers)	0	0	4.0
IV	Collection survey		2 (2 workers)	0.5 (2 workers)	4 (2 workers)	16

Note:

1. The time required to obtain the time series data is indicated for the stationary camera survey (Method I and II), whereas the time required to collect the data at a single point is indicated for the manual survey (Method III and IV). The types of data obtained from each method differ (See Table 4).
2. The time required for procuring materials for each survey, for the preliminary manufacturing of metal fittings to install the solar panel and the wiring to the camera (see Table 13), and for camera option settings is not included.
3. "Data processing" above does not include the time required for automatic processing and downloading etc., only manual processing time. Also, "Data processing" of Method IV refers to time taken to input and check the data on the PC from the on-site survey result.
4. If each worker performed a different task, describe it as 1 worker.
5. In one case, it took about two months and 15 people to prepare a beach litter dataset of 3,500 images for semantic segmentation, classifying man-made and natural litter pixel by pixel. Therefore, it is practical to use existing publicly available data.
6. Although no maintenance was conducted on the stationary camera during this demonstration, it is thought that maintenance will be necessary if the cameras are to be installed for a long period of time.

Table 8. Time spent on survey (Time required for the second and subsequent surveys)

<table>
<tr><th colspan="2" rowspan="2">Survey Methods</th><th colspan="2">Data Sampling (Unit: Hour)</th><th colspan="2">Data Processing/Analysis (Unit: Hour)</th><th rowspan="2">Total Person-Hour (Unit: Person*Hour)</th></tr>
<tr><th>Preparation</th><th>Measurement /Sampling</th><th>Processing</th><th>Detection Classification Quantification</th></tr>
<tr><td>I</td><td>Stationary camera and AI</td><td rowspan="2">0</td><td rowspan="2">0</td><td rowspan="2">0</td><td>0</td><td>0</td></tr>
<tr><td>II</td><td>Stationary camera and manual detection</td><td>0.75
(1 worker)</td><td>0.75</td></tr>
<tr><td>III</td><td>On-site visual inspection</td><td rowspan="2">0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>IV</td><td>Collection survey</td><td>2
(2 workers)</td><td>0.5
(2 workers)</td><td>4
(2 workers)</td><td>16</td></tr>
</table>

Note:

1. The time required to obtain the time series data is indicated for the stationary camera survey (Method I and II), whereas the time required to collect the data at a single point is indicated for the manual survey (Method III and IV). The types of data obtained from each method differ (See Table 4).
2. The time required for procuring materials for each survey, for the preliminary manufacturing of metal fittings to install the solar panel and the wiring to the camera (see Table 13), and for camera option settings is not included.
3. "Data processing" above does not include the time required for automatic processing and downloading etc., only manual processing time. Also, "Data processing" of Method IV refers to time taken to input and check the data on the PC from the on-site survey result.
4. If each worker performed a different task, describe it as 1 worker.
5. In one case, it took about two months and 15 people to prepare a beach litter dataset of 3,500 images for semantic segmentation, classifying man-made and natural litter pixel by pixel. Therefore, it is practical to use existing publicly available data.
6. Although no maintenance was conducted on the stationary camera during this demonstration, it is thought that maintenance will be necessary if the cameras are to be installed for a long period of time.

3.4 Comparison of survey cost

Table 9. Survey cost (Cost required for the first survey) (Unit: USD)

<table>
<tr><th colspan="2">Survey Methods</th><th>Equipment Including Tools</th><th>SIM Card and Server License Fee (/Year)</th><th>Equipment Transportation</th><th>Travel Expenses</th><th>Labor Cost</th><th>Total</th></tr>
<tr><td>I</td><td>Stationary camera and AI</td><td rowspan="2">1,822</td><td rowspan="2">340</td><td rowspan="2">55</td><td rowspan="2">646</td><td>314</td><td>3,178</td></tr>
<tr><td>II</td><td>Stationary camera and manual detection</td><td>344</td><td>3,207</td></tr>
<tr><td>III</td><td>On-site visual inspection</td><td>593</td><td>-</td><td>-</td><td>646</td><td>157</td><td>1,396</td></tr>
<tr><td>IV</td><td>Collection survey</td><td>831</td><td>-</td><td>28</td><td>646</td><td>1,587</td><td>3,092</td></tr>
</table>

Note:

1. USD = 147.10 yen (To convert yen to USD, we used the table of exchange rates for yen (TTM) on September 3, 2024. Decimal values resulting from the conversion were rounded off.)
2. The cost of setting up the equipment will vary depending on the type of camera, whether solar panels are installed, the server, SIM card and server license fee, etc. If you outsource the equipment installation to a construction company, you will also need an outsourcing fee.
3. Operating cost excludes daily allowance and accommodation costs.
4. Only travel expenses within Toyama Prefecture are included.
5. The labor cost was calculated based on the unit cost of a survey engineer applied by the Japanese Ministry of Land, Infrastructure, Transport and Tourism in 2024. The labor cost per hour is approximately 39 USD. This labor cost does not include overhead costs.
6. Although the disposal cost in this test was covered by Imizu city (a local government where the camera was installed), companies and other organizations will usually need to commission industrial waste disposal companies when conducting collection surveys (Method IV).
7. Although no maintenance was conducted on the stationary camera during this demonstration, it is thought that maintenance costs will be necessary if the cameras are to be installed for a long period of time.

Table 10. Survey cost (Cost required for the second and subsequent surveys) (Unit: USD)

<table>
<tr><th colspan="2">Survey Methods</th><th>Equipment Including Tools</th><th>SIM Card and Server License Fee (/Year)</th><th>Equipment Transportation</th><th>Travel Expenses</th><th>Labor Cost</th><th>Total</th></tr>
<tr><td>I</td><td>Stationary camera and AI</td><td rowspan="2">0</td><td rowspan="2">0</td><td rowspan="2">0</td><td rowspan="2">0</td><td>0</td><td>0</td></tr>
<tr><td>II</td><td>Stationary camera and manual detection</td><td>29</td><td>29</td></tr>
<tr><td>III</td><td>On-site visual inspection</td><td>593</td><td>-</td><td>-</td><td>646</td><td>157</td><td>1,396</td></tr>
<tr><td>IV</td><td>Collection survey</td><td>831</td><td>-</td><td>28</td><td>646</td><td>1,587</td><td>3,092</td></tr>
</table>

Note:

1. USD = 147.10 yen (To convert yen to USD, we used the table of exchange rates for yen (TTM) on September 3, 2024. Decimal values resulting from the conversion were rounded off.)
2. The cost of setting up the equipment will vary depending on the type of camera, whether solar panels are installed, the server, SIM card and server license fee, etc. If you outsource the equipment installation to a construction company, you will also need an outsourcing fee.
3. Operating cost excludes daily allowance and accommodation costs.
4. Only travel expenses within Toyama Prefecture are included.
5. The labor cost was calculated based on the unit cost of a survey engineer applied by the Japanese Ministry of Land, Infrastructure, Transport and Tourism in 2024. The labor cost per hour is approximately 39 USD. This labor cost does not include overhead costs.
6. Although the disposal cost in this test was covered by Imizu city (a local government where the camera was installed), companies and other organizations will usually need to commission industrial waste disposal companies when conducting collection surveys (Method IV).
7. Although no maintenance was conducted on the stationary camera during this demonstration, it is thought that maintenance costs will be necessary if the cameras are to be installed for a long period of time.

4. Conclusion

We demonstrated that data sampling of beach litter using a stationary camera according to the guide-lines
worked well. Matters that the stationary camera and AI for monitoring beach litter can achieve for the beach litter survey and their technical limitations that were clarified through the demonstration test are shown in Table 11. The points to be noted were extracted at each stage of the planning, preparation, and implementation of the demonstration test, and these were reflected in the Annex of the guidelines (see Table 12).
As a result of comparing the time and cost required for each method, the following characteristics were confirmed:

- Comparison of surveys using stationary cameras and manual surveys
  - Compared to the collection survey (ground truth), it was confirmed that stationary camera surveys underestimated the number of items. However, stationary cameras can significantly reduce the time and cost required for surveys, as once they have been set up, it is possible to obtain data from the second time onwards without spending time and cost. The stationary camera method is thought to be extremely useful, especially when continuously monitoring the situation.
- Comparison of surveys using stationary cameras and AI (Method I) and stationary cameras and manual detection (Method II)
  - The number of pixels detected in Method I and those of Method II showed a similar trend, suggesting that detection by AI produces results similar to those of the human eye, and that AI can be used as a substitute for human detection.
  - As AI detection (Method I) does not require time and cost for analysis, it is particularly effective for analyzing large amounts of data and obtaining time series data. Although the AI used in this demonstration does not classify litter or estimate the number of items, there are also AIs that can perform those tasks. As a future step, it is expected that it will be possible to classify litter and estimate the number of items using AI from images taken by stationary cameras.
  - On the other hand, manual detection (Method II) can classify litter and estimate the number of items, and it can be used even when the AI is not available.

Table 11. Matters that stationary camera and AI can achieve for beach litter survey

What Stationary Camera and AI can do for Beach Litter Survey	Current Technology Limitations
Quantification of amount of beach litter (pixels) and acquisition of time series data automatically using semantic segmentation	− If smaller beach litter is deposited at low height and low density, the error will be greater. On the other hand, when litter is densely drifted ashore, it is difficult to separate and quantify the different types of litter for the reasons described above. − Automatic analysis of the piece count is generally not suitable for semantic segmentation, as it recognizes overlapping objects as a single piece of litter, which can lead to large errors.
Obtaining the number of pixels, piece count, and the litter covered area of a detectable size man-made object.	− There is a minimum size of litter that can be detected by stationary cameras and AI. − It is difficult even for manual detection of beach litter from images to identify litter from the background when the resolution is small, and the classification accuracy is not as high as that of human investigation. − Although the number of pixels occupied by litter can be obtained from the stationary camera image, it should be noted that the number of pixels and the litter covered area may not be in a proportional relationship if the amount of litter differs between the front and back as the image is taken from an oblique angle. − There is also a possibility that it will be difficult to detect and classify litter that is in the back of the image taken by stationary cameras, as it will appear small in the image.

Table 12. Several items were identified that should be added to the Annex of the guidelines

1 Selection of Survey Points
− If there is vegetation directly above the target, it may not be possible to photograph the target from stationary cameras. − Cooperation from the local governments and beach managers can be obtained. − No cleaning activities are conducted periodically, or cleaning is conducted but litter tends to accumulate in a short period of time, making it easy to observe trends in litter abundance. − The beach is not used much (as it can spoil the landscape, and the presence of people can be an obstacle to image analysis). − Locations where there is not a lot of vegetation or snowfall (they may hinder image analysis). − Locations that are representative of the entire beach, based on the characteristics of the beach in terms of the amount and type of litter.
2 Installation of the Stationary Camera
− The closer the linear distance from stationary cameras to litter, the greater its resolution, and the closer the angle is vertically from the ground, the easier it is to identify the shape of the litter. Taking these considerations into account, it is advisable to adjust the altitude and angle of stationary cameras. − When deforming structures of beaches for installation, such as drilling a hole in the concrete of an embankment, it is necessary to confirm with the local governments and beach managers, including how to restore the site at the end of the survey.
3 Safety Management
− It is imperative that bolts and other objects do not remain after they are removed to ensure safety. − Since the equipment needs to be installed using tools, trained personnel should be engaged in the work wearing protective equipment. If it is difficult to perform the work directly, a construction company may be asked to design and install the equipment.

4	Time Required for Survey and Analysis
−	Although stationary cameras require more time to obtain data than manual surveys, once installed, time series data can be obtained even remotely, which is difficult to obtain with manual surveys. Method I (stationary camera and AI) requires almost no time to obtain litter data compared to collection surveys because litter detection is performed automatically. Considering the above, the cost-effectiveness of analysis using stationary cameras is considered to be significant.
−	Although the AI sometimes mistakenly detected sun glint or fog of the camera lens, in the case of time series data, it is thought that the effects of these abnormal data can be reduced by taking a moving average over several days. In this demonstration, a 7-day moving average was used, but the appropriate number of days for averaging may differ depending on the characteristics of regions, such as the frequency of rainfall.
−	The litter in the front appears larger and the litter in the back appears smaller in a photograph, and there is a difference in the number of pixels. In order to represent the amount of litter more accurately, the litter covered area should be calculated by projection transformation. However, instead of the litter covered area, it is thought that the approximate trend of the temporal variation in the amount of litter can be grasped from the number of pixels of litter (before the projection transformation). It is still important to note that if there is a difference in the amount of litter between the front and back of the image, there is a possibility that the amount of litter and the number of pixels will not be in a proportional relationship.

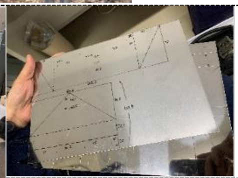
Appendix B

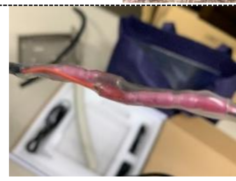
Detailed information regarding each survey methods

I Camera installation

A hole was drilled in the concrete of the embankment, and a single pipe was fixed with a metal saddle band, and a stationary camera was installed.
Rokudoji Beach was chosen as the survey site as the local community association, which has been involved in cleanup activities for many years, had provided information that the local litter tends to accumulate on the beach.
The beach litter in the survey site (Rokudoji beach) is characterized by the large amount of litter that originates from land upstream, because it is located at the mouth of the Oyabe River and Shou River. In addition, litter is expected to accumulate particularly around the base of the breakwater near the embankment. Regarding these characteristics, the camera was installed inside the breakwater near Oyabe River to understand the drift of trash that originates from land (see Figure 2).
The camera shooting coverage, camera altitude, and angle were set as shown in Figure 3 and Figure 4, regarding the resolution of the litter, and the operability of the camera. In addition, a solar panel was installed to obtain a continuous power supply during the survey. The metal fittings to install the solar panel and the wiring to the camera were manufactured and prepared in advance (see Table 13).

Table 13. Manufacturing of metal fittings to install the solar panel and the wiring to the camera

No.	Procedure	Photograph
1	Manufacturing of steel plate that serves as the base for solar panels	
	– 1 mm thick steel sheet was cut with a shearing machine. – The burr (the part of the workpiece that protrudes from the edge due to shearing, etc.) was removed.	 
	– Holes were made in the steel plate for screwing with a drilling machine.	 
	– The steel plate was bent according to the design using a bending machine. The angle of the bend was determined regarding the structure of the solar panel.	 
	– The steel plate was chamfered using a grinder, and then attached to the solar panels.	 

2	Modification of the wiring plug to connect to the camera	
	− For the wiring from the solar panel to the camera, the existing straight plug was replaced by an L-shaped plug with higher stability. − The existing plugs were stripped of their insulation, the positive and negative wires of the L-shaped plugs were connected using tools, and the wires were covered with plastic tubing for further protection.	 
	− The holes in the case (made of steel) were enlarged by drilling because the wiring plug was changed from a straight type to an L-shaped plug (described above), which will not fit inside an existing camera case.	 

II Stationary camera survey

Table 14. Metadata of stationary camera survey

<table>
<tr><td rowspan="5">Survey plan and preparation</td><td>Survey period</td><td>29th June 2024 ~ 31st December 2024</td></tr>
<tr><td>Survey date of on-site visual inspection and collection survey</td><td>3rd September 2024 1 p.m.
The weather is sunny.
(Low tide time : 8 a.m.)</td></tr>
<tr><td>Survey target</td><td>Man-made objects over 2.5 cm
In Survey I, only the total amount of man-made objects is obtained since the AI cannot perform detailed waste sorting.
The types of items to be surveyed in Survey II, III, and IV are plastic bottles, food containers (incl. fast food cups, etc.), plastic bags (shopping bags, food bags, industrial bags, etc.), seedling pots, and other man-made litter, based on the characteristics confirmed in the preliminary survey.</td></tr>
<tr><td>Regulations and legislation</td><td>Permission to install stationary cameras for local government and beach managers</td></tr>
<tr><td>Roles of researcher</td><td>Preparation (including manufacturing of metal fittings): 1 person
Camera installation / removal: 1 person
Litter visual inspection: 2 people
Litter collection survey: 2 people</td></tr>
<tr><td rowspan="5">Camera setting</td><td>Camera</td><td>HykeCam LT4G (Hyke, Inc.)</td></tr>
<tr><td>Shooting duration</td><td>AM 5:00 – PM 7:00 every day, photographed once an hour</td></tr>
<tr><td>Number of pixels per image</td><td>4,032 x 3,024 pix</td></tr>
<tr><td>Camera installation altitude</td><td>2.7 m above the ground at the beach</td></tr>
<tr><td>Camera angle</td><td>+70°</td></tr>
<tr><td rowspan="2">Survey implementation</td><td>Installation of survey equipment (grand control point)</td><td>Deployment of ground control points (GCPs) is described in Figure 18. A total station was used for measurement.</td></tr>
<tr><td>Survey area (photographing area)</td><td>70.933 m^2</td></tr>
<tr><td>Image processing</td><td>I. AI detection</td><td>To estimate the litter covered area, positional information correction was conducted on the taken image by projection transformation using surveying technology.</td></tr>
<tr><td rowspan="2">Image analysis</td><td>I. AI detection</td><td>AI developed by Dr. Kako (Kagoshima university) was used for the beach litter detection.</td></tr>
<tr><td>II. Manual detection</td><td>Annotation software (Labelme (the Massachusetts Institute of Technology (MIT)) was used to visually surround areas of litter from the images, and the number of pixels in these areas was counted using Python.</td></tr>
</table>

Figure 18. GCPs (ground control points) of the survey site

Figure 19. Surrounding survey area by rope

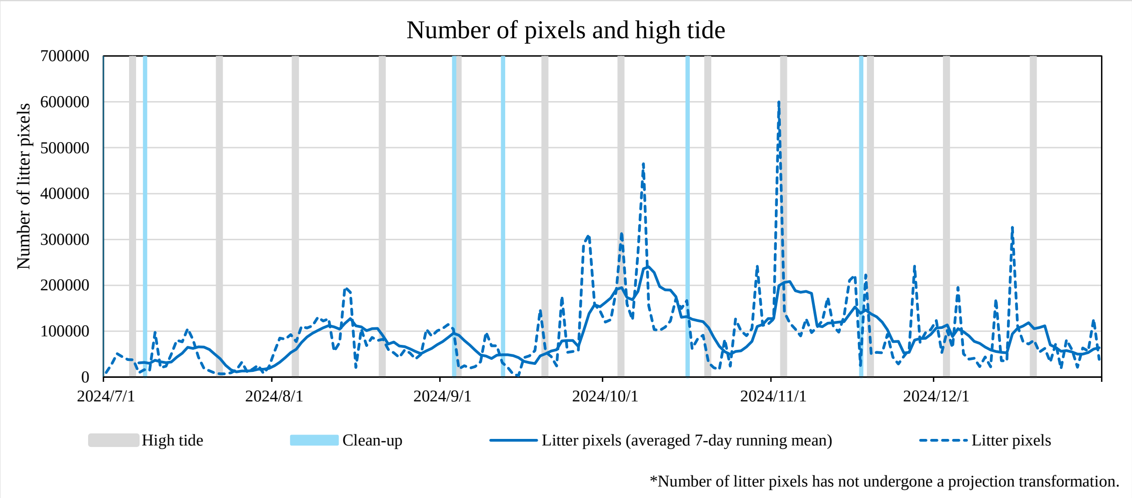


Figure 20. Time series data: Number of pixels and duration of high tide

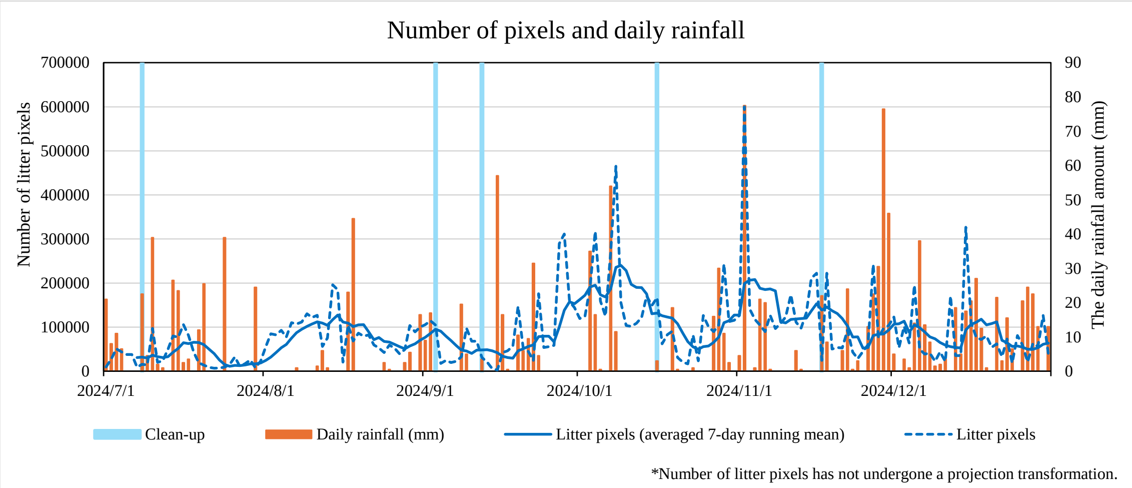


Figure 21. Time series data: Number of pixels and rainfall

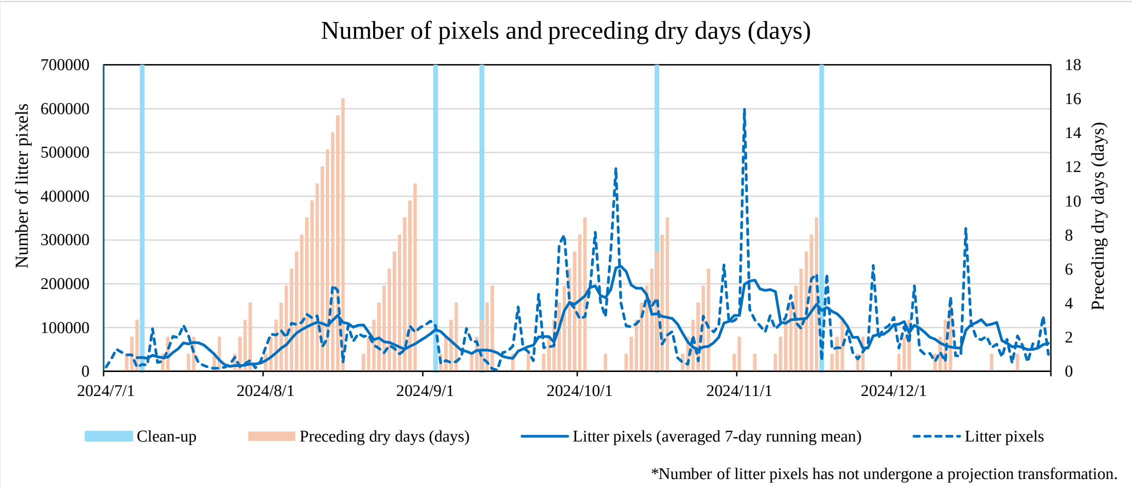


Figure 22. Time series data: Number of pixels and preceding dry days

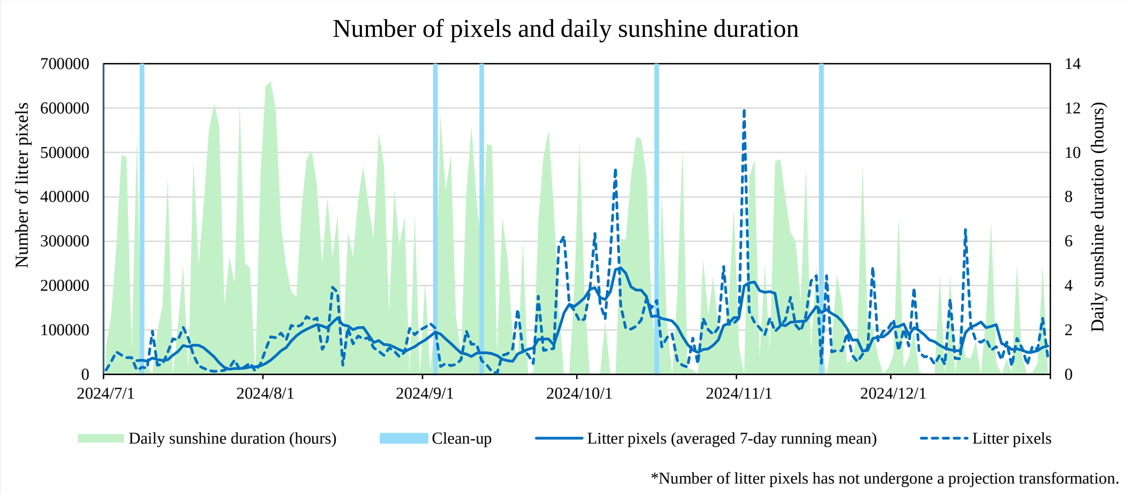


Figure 23. Time series data: Number of pixels and daily sunshine duration

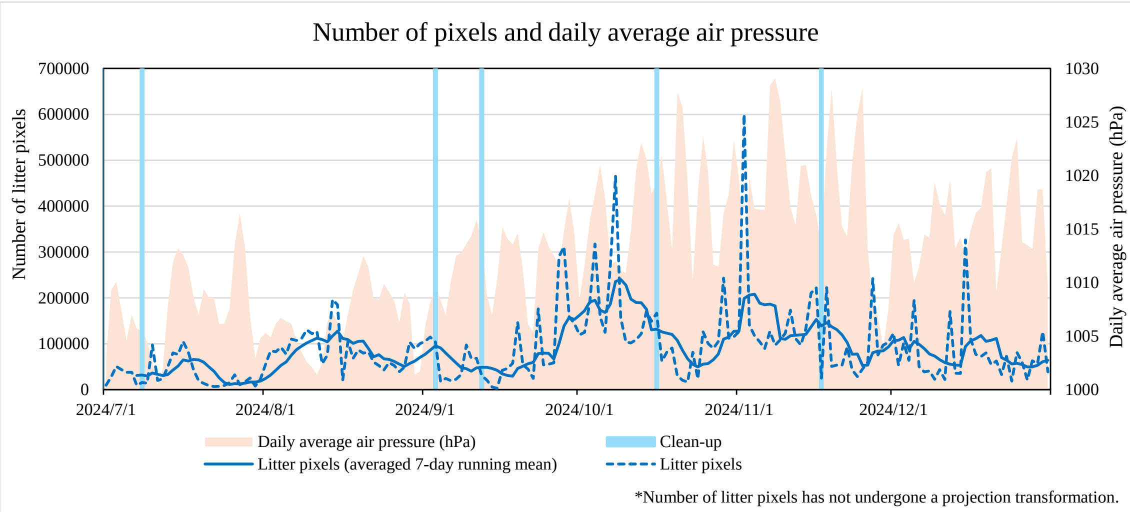


Figure 24. Time series data: Number of pixels and air pressure

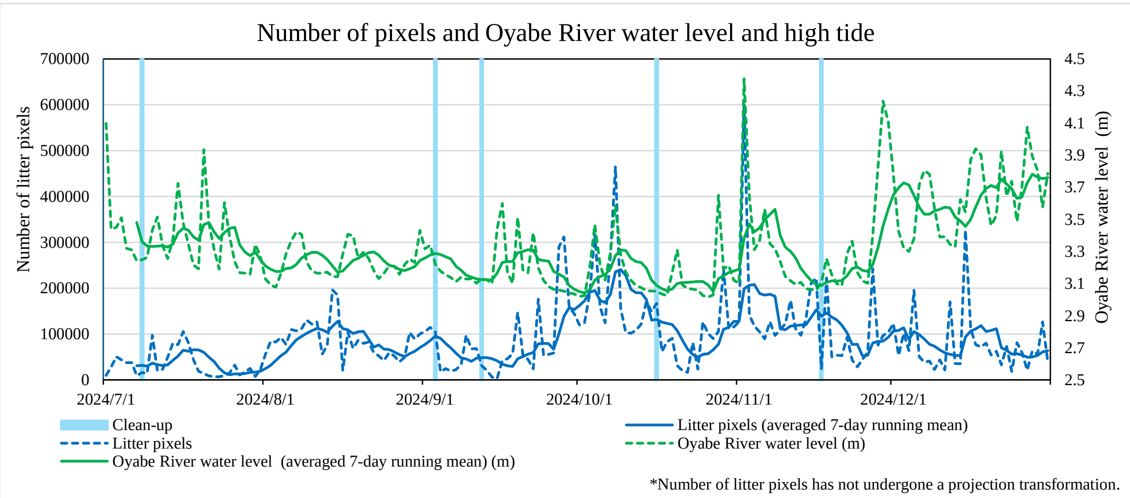


Figure 25. Time series data: Number of pixels and Oyabe River water level

Table 15. Results of piece count comparison between stationary camera and manual detection and the ground truth (in Survey plot 1)

Objects	Stationary Camera and Manual Detection (Number)			Ground Truth (Number)	Detection Rate (Average of Manual Detection / Ground Truth)
	Researcher α	Researcher β	Average		
Plastic bottles	4	4	4	4	1.00
Food containers (incl. fast food cups, etc.)	5	4	5	8	0.56
Plastic bags (for shopping, food, business use, etc.)	0	0	0	0	0.00
Seedling pots	1	0	1	1	0.50
Other man-made objects	54	42	48	52	0.92
Total	64	50	57	65	0.88

Note:
1. Estimations of the stationary camera and manual detection are the average of the survey results from each of the two researchers.
2. See II and IV in Appendix B for the methods to conduct surveys and estimates.

Table 16. **Results of piece count comparison between stationary camera and manual detection and the ground truth (in Survey plot 2)**

Objects	Stationary Camera and Manual Detection (Number)			Ground Truth (Number)	Detection Rate (Average of Manual Detection / Ground Truth)
	Researcher α	Researcher β	Average		
Plastic bottles	4	2	3	4	0.75
Food containers (incl. fast food cups, etc.)	1	1	1	21	0.05
Plastic bags (for shopping, food, business use, etc.)	0	0	0	0	0.00
Seedling pots	2	1	2	5	0.30
Other man-made objects	106	84	95	155	0.61
Total	113	88	101	185	0.54

Note:
1. Estimations of the stationary camera and manual detection are the average of the survey results from each of the two researchers.
2. See II and IV in Appendix B for the methods to conduct surveys and estimates.

Table 17. Results of piece count comparison between stationary camera and manual detection and the ground truth (in Survey plot 3)

Objects	Stationary Camera and Manual Detection (Number)			Ground Truth (Number)	Detection Rate (Average of Manual Detection / Ground Truth)
	Researcher α	Researcher β	Average		
Plastic bottles	4	4	4	9	0.44
Food containers (incl. fast food cups, etc.)	1	1	1	16	0.06
Plastic bags (for shopping, food, business use, etc.)	0	0	0	0	0.00
Seedling pots	2	1	2	2	0.75
Other man-made objects	38	38	38	99	0.38
Total	45	44	45	126	0.35

Note:
1. Estimations of the stationary camera and manual detection are the average of the survey results from each of the two researchers.
2. See II and IV in Appendix B for the methods to conduct surveys and estimates.

Table 18. Results of piece count comparison between stationary camera and manual detection and the ground truth (in Survey plot 4)

Objects	Stationary Camera and Manual Detection (Number)			Ground Truth (Number)	Detection Rate (Average of Manual Detection / Ground Truth)
	Researcher α	Researcher β	Average		
Plastic bottles	9	8	9	18	0.47
Food containers (incl. fast food cups, etc.)	2	0	1	29	0.03
Plastic bags (for shopping, food, business use, etc.)	0	0	0	1	0.00
Seedling pots	1	1	1	6	0.17
Other man-made objects	46	45	46	184	0.25
Total	58	54	56	238	0.24

Note:
1. Estimations of the stationary camera and manual detection are the average of the survey results from each of the two researchers.
2. See II and IV in Appendix B for the methods to conduct surveys and estimates.

III On-site visual inspection

(1) Calibration

In order to reduce errors of the on-site visual inspection at the actual measurement site (Method III), a preliminary litter evaluation test (comparing the results of the visual estimation of volume with the actual volume of the litter) was conducted at two other sites (see Table 19).

Table 19. Result of calibration

Location	Researcher	On-site Visual Inspection (L)	Ground Truth (L)
Site A	α	3	20
	β	5	
Site B	α	11	50
	β	45	

(2) Comparison of volume estimation results (Method II and III)

We evaluated the volume of beach litter as "the number of 20 L bags filled with litter within 50 m of the shoreline at 5 m intervals".

Table 20. Comparison of the results of volume estimation using stationary cameras and manual detection (Method II), on-site visual inspection (Method III), and ground truth (Unit: m^3)

Objects	Stationary Camera and AI			On-site Visual Inspection			Ground Truth
	Researcher α	Researcher β	Average	Researcher γ	Researcher δ	Average	
Total litter	0.040 (0.024-0.048)	0.031 (0.024-0.048)	0.035 (0.024-0.048)	0.14	0.10	0.12	0.074

Note:
1. Values are rounded to the fourth decimal place.
2. Values in parentheses indicate a range of estimated quantities using a rank table (see Table 20), taking into account the uncertainty involved in estimating the volume from oblique-angle photographs. The estimated values were obtained by multiplying the values by 1.2, as the average length of the shoreline of the surveyed beach was approximately 12 m (see Table 21 for the estimation method).

Table 21. Rank table for Method II

Rank for Volume Estimation	Number of 20L Garbage Bags	Estimated Volume (m^3/10m)
0	0	0
1	0~1	0~0.02
2	1~8	0.02~0.16
3	8~64	0.16~1.28
4	64~	1.28~ (*)

* Based on the same report, the maximum volume was set at 2.2 m^3.
Source: Beach litter spatial distribution survey report (2023, Ehime prefectural government and JANUS)

IV Collection survey

(1) Collecting beach litter
Man-made objects over 2.5 cm and driftwood were collected manually. In the case where such large marine litter that cannot be collected by human power is found in the investigation area, measure the dimensions of the marine litter to estimate the volume.

(2) Classification of beach litter
The classification item list used in the UAV demonstration test for beach litter (see Appendix 1) was reorganized based on the characteristics of the survey site for this demonstration test (Table 22).

(3) Measurement for quantification
The number, volume, and weight of each item were measured (See Table 23). The pictures of each item were taken. And the number, volume, and weight of each item were recorded in the data sheet with meta data (e.g., survey date, matrix of shore, survey point coordinates).

Table 22. Classification list for demonstration test

material	item type	
plastic	bottle cap and lid	
	plastic bottle	for drink
		other
	straw	
	cutlery	
	food containers incl. fast food cup etc.	
	plastic bag	
	lighter	
	tape incl. package band	
	pieces of sheet/bag	
	pieces of hard plastic	
	sponge	
	buoy (fishing gear)	
	rope, string (fishing gear)	
	conger tube (lid, cylinder)	
	small pipe for oyster culture (1.5cm in length)	
	pipe for oyster culture (10 to 20cm in length)	
	fishing net	
	other fishing gear	
	fishing tackle (for recreation)	
	cigarette butt/filter	
	household goods incl. tooth brush	
	seedling pot	
	other plastic	
foamed plastic	float/buoy made of foamed polystyrene	
	pieces of foamed polystyrene	
	other foamed polystyrene	
rubber	rubber	
glass, ceramic	glass, ceramic	
metal	metal	
paper, cardboard	paper, cardboard	
fabric	fabric	
other	other	
large marine litter that cannot be collected by man power	coordinates item type ( )	

Table 23. Result of collection survey

material	item type		number	volume (L)	weight (kg)
plastic	bottle cap and lid		58	1.15	0.181
	plastic bottle	for drink	26	11.3	1.644
		other	9	6.3	0.336
	straw		5	0.021	0.006
	cutlery		0	0	0
	food containers incl. fast food cup etc.		74	10.87	0.697
	plastic bag		1	0.4	0.029
	lighter		10	0.39	0.13
	tape incl. package band		31	2.01	0.045
	pieces of sheet/bag			4.2	0.171
	pieces of hard plastic			14.1	1.289
	sponge		31	2.1	0.234
	buoy (fishing gear)		0	0	0
	rope, string (fishing gear)		0	0	0
	conger tube (lid, cylinder)		0	0	0
	small pipe for oyster culture (1.5cm in length)		0	0	0
	pipe for oyster culture (10 to 20cm in length)		0	0	0
	fishing net		0	0	0
	other fishing gear		0	0	0
	fishing tackle (for recreation)		0	0	0
	cigarette butt/filter		2	0.001	0.002
	household goods incl. tooth brush		8	0.68	0.063
	seedling pot		14	3.1	0.098
	other plastic		107	19.3	0.735
foamed plastic	float/buoy made of foamed polystyrene		0	0	0
	pieces of foamed polystyrene			6.05	0.915
	other foamed polystyrene		218	7.5	0.242
rubber	rubber		4	0.037	0.07
glass, ceramic	glass, ceramic		2	0.3	0.167
metal	metal		1	0.4	0.035
paper, cardboard	paper, cardboard		1	0.001	0.001
fabric	fabric		0	0	0
other	other		12	0.85	1.047
large marine litter that cannot be collected by man power	coordinates item type (　　　　　)		0	0	0
		Total	614	91.06	8.137