Technical Report

Counting entering strollers with AI and analyzing patterns at a sightseeing spot

Ryo Hori (Graduate School of Informatics, Nagoya University, hori.ryo@f.mbox.nagoya-u.ac.jp)
Hikaru Misaki (School of Informatics, Nagoya University, misaki.hikaru@g.mbox.nagoya-u.ac.jp)
Tomoki Uga (School of Informatics, Nagoya University, uga.tomoki@f.mbox.nagoya-u.ac.jp)
Mayu Urata (Graduate School of Informatics, Nagoya University, mayu@i.nagoya-u.ac.jp)
Mamoru Endo (Graduate School of Informatics, Nagoya University, endo@i.nagoya-u.ac.jp)
Takami Yasuda (Graduate School of Informatics, Nagoya University, yasuda@i.nagoya-u.ac.jp)

Abstract
Nowadays, it is thought important to make a decision based on data to use limited resources effectively. Strollers, one of the main targets of sightseeing places for children, are believed to be important for such places, but ticket sales or technology recently utilized are not suitable to grasp the movements of the strollers, therefore, this research focused on strollers through the use of a camera and artificial intelligence. We, first, recorded an entrance, and fine-tuned prepared models to create original object detection artificial intelligence. Then, we counted the number of entering strollers, and analyzed the entering patterns. We discovered several patterns, and suggest that there is a difference between strollers in the weekdays and weekends. Lastly, for future prospects, we mention two things, a real-time counting system with GPU installed small computers, and simultaneous counting of strollers' companions.

Keywords
tourism, EBPM, object counting, data utilization, zoo

1. Introduction

Recently in Japan, with the population decreasing and the society aging, the workforce is shrinking. In order to deal with such a situation, people are trying to make their work more efficient by utilizing artificial intelligence (hereinafter called "AI"). Indeed, the Japanese post office has started to answer redelivery calls with AI, and websites of hotels automatically answer questions about staying by guests. AI which performs simple tasks instead of humans is spreading widely.

In addition, Evidence-Based Policy Making (hereinafter called "EBPM") is considered to be important in every field [Somusho, 2020]. Analyzing the current situation or measuring effects of measures with the basis of data helps to use a limited budget or human resources effectively.

In this paper, we report a practice of collecting data in the real world by using AI and analyzing it for EBPM.

2. Situations

We, Nagoya University, applied with NEC Solution Innovators, Ltd. for Hatch Technology Nagoya, a project which promotes research and development or social implementation of technologies.

Higashiyama Zoo and Botanical Gardens, a sightseeing park which Nagoya City in Japan manages, would like to invite more people and make themselves more attractive based on data. They think that it is necessary to do three things. First of all, find possible visitors or determine when to take what measures to increase the number of visitors by acquiring their nature comprehensively and continuously. Secondly, because the majority of frequent visitors are from three neighboring prefectures, they want to attract visitors from other regions. Last, take questionnaire surveys or hold events for visitors whose needs of the park have not been grasped, in order to understand them.

Now, however, some pieces of visitor's information are not acquired for some reasons: Tickets are sold through many channels such as convenience stores, travel agencies, company's welfares, and etc. In addition to this, the admission fee is free for junior high school students and under, requiring no tickets. Thus, it is difficult to obtain visitors' information through ticket sales and grasp the whole visitors' nature. Although there exist staff at entrances, they have tasks to do other than observing visitors, so they cannot concentrate on counting all the time or observe the full attributes of them. It is difficult for the park to obtain continuous information through a year.

Furthermore, participants of the surveys are various, so there will be some visitors' needs which have not been recognized.

As we showed above, the park has trouble collecting visitor's data, and requires a way to do it.

Recently, according to Morio et al. [2018], we can obtain information of people moving through mobile spatial statistics, GPS data, or Wi-Fi packet sensor data. Mobile spatial statistics is demographics provided by NTT DOCOMO, Inc. with about 24 hour 80,000,000 user samples [NTT DOCOMO, n.a.]. The data contains foreign tourists as well as Japanese. Nakanishi et al. [2018] gathered origin-destination data by Wi-Fi packet sensors placed in the Okinawa area. Murai et al. [2020] also used them to analyze behavior patterns in large-scaled leisure facilities. There is much research using Wi-Fi packet sensors [Arreeras et al., 2019; Tanaka et al., 2019]. Shingai et al. [2016] is researching the utilization possibility of mobile spatial statistics. Hashimoto et al. [2017] used the statistics to analyze effects of a new super express. These ways are valuable in that they can capture not only Japanese tourists, but also inbound ones for 24 hours, and they can observe comparatively extensive, city to city or prefecture to prefecture, movements. In addition, Matsuyama City in Ehime Prefecture and NTT DOCOMO, Inc. conducted a demonstration experiment with the cooperation of foreign tourists in the city, where they distributed Bluetooth Low Energy tags to the tourists and, with the collected data, analyzed tourists' migratory behaviors [Matsuyama City, 2019].

However, because not all people have a smartphone, or have it with Wi-Fi, or some people have a few devices, the data acquired by Wi-Fi packet sensors or mobile spatial statistics is partial or just an estimated value. Getting overall information is difficult with them. The experiment using Bluetooth Low Energy tags was still under a prepared condition.

On top of that, the ways are not suitable and other methods are required because infants or children, one of the main targets especially in a park, do not possess such devices now.

Thus, we used a camera and AI. There is research about counting systems and products using them. For example, we can name OpenDataCam, a useful open source object counting tool, which never records any photo or video data. However, it uses counting lines, requiring a good quality of object tracking. In our research, sometimes hiding behind their users, strollers, the object on which we focus, are not always captured clearly unlike cars in its demo movie. Thus, we prepared our original system which does not need strict detection or tracking, but would work well in our environment. Furthermore, social implementation is not commonly spread in the society. We value not only creating a counting AI but also applying a system into the park, or the real world more generally.

In this context, as a research of Hatch Technology Nagoya, we collected data using a camera and AI, and analyzed it at Higashiyama Zoo and Botanical Gardens.

3. Research objectives

Based on the situation above, as a practical example of collecting and analyzing data with AI, we focused on strollers, for the park does not have any ways of collecting data, which cannot be obtained through entrance ticket sales or membership registrations.

We believe sightseeing places should put an emphasis on strollers: "Iko-yo", one of the largest outings for kids information websites, has such a filtering condition as "strollers OK" [Iko-yo, n.a.]. In addition, in a questionnaire towards families with children, when they were asked what they think is important at their destinations and accommodations when traveling with children, it ranked 1st whether or not they, families with kids, are welcomed [Cozre, 2019]. There are many obstacles for strollers such as stairs, slopes, steps, etc at tourist spots, and strollers are difficult to move with. Therefore, comfortableness for strollers, care or movement lines for them, can be a key factor when deciding where to visit. If the ratio of families with strollers is revealed, it can help assess effectiveness of questionnaire surveys or events for stroller users. Besides if we know when many strollers visit the park, holding effective surveys or events at the very time will become easier, and then the park will be able to appeal as a stroller-friendly place. These are the reasons why we focused on strollers and analyzed their entering patterns.

In this research, we aimed at the two following things: counting entering strollers using AI, and analyzing the number by the hour at an entrance of Higashiyama Zoo and Botanical Gardens.

4. Counting entering strollers using YOLO

We counted them at an entrance utilizing YOLOv5 (hereinafter called "YOLO") [Redmon et al., 2016] and an original algorithm. We selected YOLO, a fast object detection AI model, taking in consideration real time detection in the near future. We recorded the entrance, trained models with the data, and had the models infer the records and output whether a stroller is in each frame into csv files based on the algorithm using a computer loaded with a GPU. The camera was installed under the entrance roof with an angle of about 30 degrees downward from parallel to the ground.

4.1 Training method

We converted the videos into pictures frame by frame, and annotated strollers in them. With this dataset, we fine-tuned default models of YOLO because originally they are not trained to recognize strollers. The recording was eight hours a day from 9 am to 5 pm from 2020/12/8 through 2020/12/13. The resolution of the videos was 1289 × 729 with 15 of frame rate. For the records of the days 8th, 9th, and 10th, we annotated fractions of strollers (about 40 %), strollers some part of which is hiding behind a person (Figure 1), as well as the whole image (about 60 %). As to one of the day 11th, we did not annotate the video to use for evaluating models. As to the days 12th and 13th, we annotated the whole image. With these annotated datasets, we trained models. We named each model that we created from A to F. Date means the day when the video was recorded from 8th to 13th; Strollers, the unique number of strollers in the videos; Train and Valid, the number of pictures to train or validate a model with; Weight, the size of YOLO weights (Table 1). For example, we used the record of day 13th to train model A with the S sized weight. 66 strollers appeared

Figure 1: A stroller behind a person (left), and the whole image well detected by AI (right)

Note: People are the authors.

Table 1: Details of datasets and results of counting

Model	Date	Strollers	Train	Valid	Weight	Precision	Recall	F1
A	13	66	1,116	579	S	21/23	21/22	0.933
B	13	66	1,116	579	X	18/19	18/22	0.878
C	12, 13	142	2,359	1,124	S	21/22	21/22	0.955
D	12, 13	142	2,359	1,124	X	19/20	19/22	0.905
E	8-10, 12, 13	204	7,953	3,607	S	22/26	22/22	0.917
F	8-10, 12, 13	204	7,953	3,607	X	21/22	21/22	0.955

in the video. The training dataset contained 1116 images, and validation datasets, 579. There were pictures in which the same strollers appeared but the directions were different.

4.2 Counting method

While YOLO is able to detect objects in every frame, it is not possible to consider the frame to frame relationship: YOLO cannot count the unique number of objects in videos. Therefore, a counting method is necessary. In this research, we counted one stroller when it was detected for certain frames in a given interval as follows. This method was possible under the condition that not so many strollers came, and they passed through the entrance gate one by one with a certain interval, and that the gate was one-way. The count variable in the following step means the number of frames where strollers were detected, the endure, not detected, and the record allows us to count the next stroller. We calculated the time of every frame based on FPS (15 fps) and the time at which the video record started. YOLO sometimes gives a bounding box to other objects such as people, clothes, or background. In order not to count them, we used a threshold of five frames. Also, it was necessary to decide whether a stroller has passed the gate. We defined more than 50 frames where no stroller was detected means that a stroller has passed through. The numbers of five or 50 were determined by our experience when we annotated the videos. Detailed method is as follows.

1. Initialize count as zero, endure as zero, record as True.
2. If record is True, and AI detects a stroller, increase count by one, and set endure to zero.
3. If count is five, set record False.
4. If AI does not detect it, increase endure by one.
5. If endure is more than 50, set record True, and count zero.
6. Output to a csv file time calculated by FPS and whether count is more than five (True or False).
7. Repeat 2 to 6 steps towards all frames.

4.3 Results of counting

We made a short video, where 22 strollers, extracted from the eight hour video of the day 11th enter the gate, so that we can test our original models with ease (Figure 1 and Table 1). Note that not all strollers in the eight hour video appear in the short one. Although there is a time in the video when no strollers or no people pass through the gate, the video does not contain time when other objects, mainly only people, do so. Therefore, precision may be higher than real. The Precision, Recall, F1 measures in Table 1 were against this video. We evaluated the whole system, AI model and count algorithm. For instance, as to the model A, precision, recall, f1 were 21/23, 21/22, 0.933 respectably toward the strollers in the video.

4.4 Discussion of counting

As to training, it was possible to make good models with only one-day or two-days dataset and an S sized model, the smallest model that YOLO prepares (see model A and C in Table 1). General datasets contain from tens of thousands of images to millions of images. Compared to this size, 1,116 to 2,359 pictures are very small. This means that we can realize counting with short time training and low equipment cost. On the other hand, strollers hiding behind people were difficult to detect, and it became harder to train model with datasets including such images. Therefore, environmental arrangements, where to install a camera and the direction to set it, were very important. We expect that by installing a camera vertically to the ground under the entrance roof, strollers will not hide behind people, so training models will be easier and the accuracy of inference will increase.

As to counting, the original algorithm worked well with our environment. There were almost no counting errors, so the accuracy depended on how well AI could detect strollers. Accordingly, for more accurate and stable counting, we need more precise detection by better training or environmental arrangements.

5. Analysis of the number of strollers by hour

Based on the csv files which we made using the system above, we visualized three things: The number of entering strollers by the hour for five days, the average number and cumulative percentage of them, for weekdays, weekends, and all days. Hour in x axis of the figures has one hour of time span: Black circled point at nine hour of the day 10th in Figure 2 means 10 strollers entered the park between 9 am to 10 am. Sum in the legends in Figure 2 and 3 means the total number of the strollers of the day or average.

5.1 Result of analysis

Figure 2 shows that at 1 pm to 2 pm there is a large difference in the number between the weekdays and weekends. Figure 3 shows that in the weekends there came 1.85 times as many strollers as in the weekdays. Figure 4 shows that by 11 am 60 % of the strollers had entered, and by 1 pm more than 80 %. Especially, in the weekdays (square pointed line in Figure 4), they had passed through the gate comparatively early. Figure 2 and Figure 3 show that more strollers came one or two hours after the park opened rather than immediately after the opening. After 3 pm, almost no strollers visited the park. Lastly, on the assumption that families with strollers are a family of four on average, 40 % of all the visitors used a stroller. We obtained the number of people from infrared counters, which the park has installed.

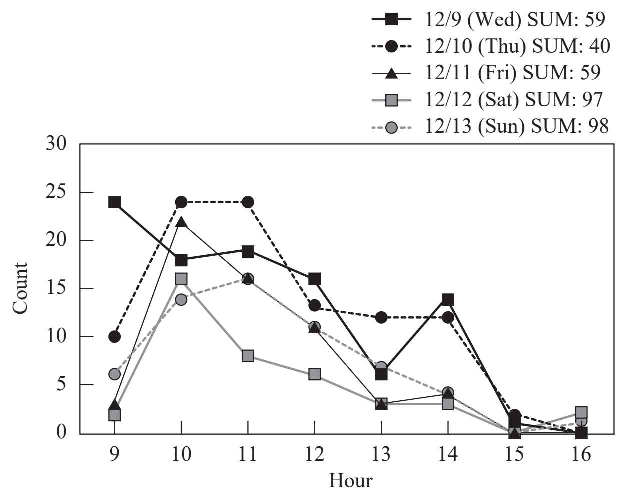

Figure 2: Number of entering strollers by hour for five days

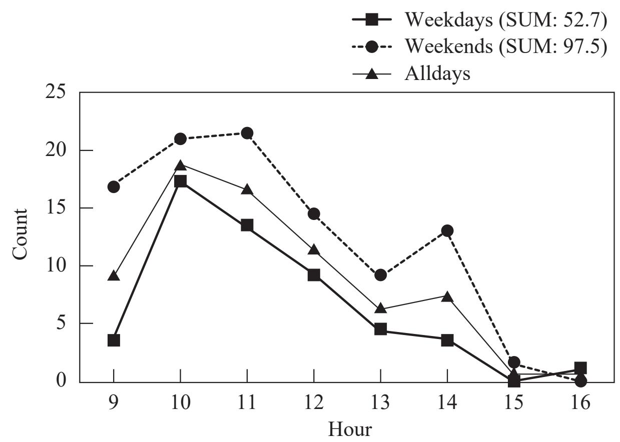

Figure 3: Average number of strollers for weekdays, weekends, and all days

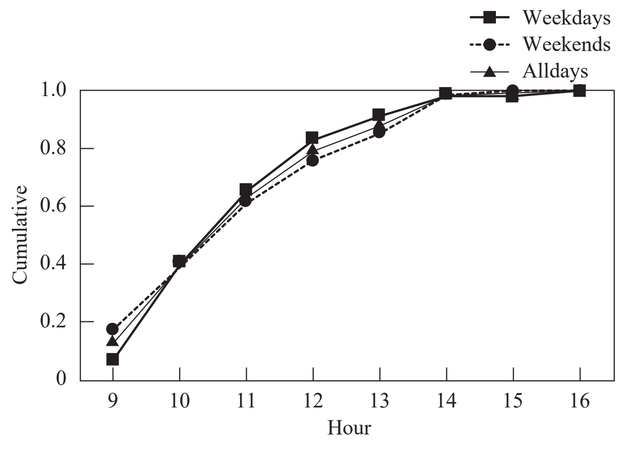

Figure 4: Cumulative percentage of strollers for weekdays, weekends, and all days

5.2 Discussion of analysis

From these results, some possible measures are as follows:

- Because many strollers entered the park before noon, conducting surveys then will be effective.
- This also means that they may have lunch in the park, so making the environment there more comfortable may raise visitors' satisfaction level.
- Preparing recommended route maps for strollers may lead to more satisfaction.

Through these measures, the park will be able to acquire frequent visitors.
In addition:

- Because about one in five strollers may take lunch outside the park, the park could receive more visitors to eat lunch inside it if some event is held before noon.
- In the evening, the ratio may decrease because 3 pm or 4 pm can be thought to be the time when families with children start preparing for dinner. Thus, it may be good to focus on visitors who do not carry a stroller or aim to acquire new visitors who have not come there by offering evening tourist plans. However, we need to compare the number of people and strollers by the hour to assert that the ratio really decreases.

Moreover, in the weekends, strollers increased 1.85 times, and human visitors, 2.48. From this fact and our experience while annotating, we suggest that strollers in weekends and weekdays are different: The former was mother and her child, and the latter was family using a stroller.

We discovered some patterns; however, it is unknown if we can observe these patterns at anytime. Counting longer term will lead to more accurate and stable patterns.

Lastly, the number we counted does not equal the number of strollers in the park. By counting that of exiting strollers, the park will know how many strollers are in it. Thus, it will become easier to determine the timing to hold events.

6. Discussion

Utilizing AI and arranging environments for it, we were able to collect data about entering strollers by the hour automatically, which information the park could not obtain conventionally, and discover some entering patterns.

Based on the data, the park will be able to determine whether or not to make investments for a more comfortable park environment. That we were able to observe strollers automatically will lead to comprehensive and continuous data collection, comparison of the effects of seasons or weather, and measuring effects of events. It will also be possible to conduct questionnaire surveys and comprehend strollers' needs effectively, or hold events which entertain more people based on the data. Through these measures, the park will be a more attractive sightseeing place.

As we collected data which could not be obtained through ticket sales, it will also be possible to understand visitors at similar open sightseeing spots such as Ai • Chikyuhaku Kinen Koen, Gifu Family Park, Outlet Malls, or Okage Yokocho. Moreover, not only at sightseeing places, but also at highway service areas or bus stops, counting certain objects is possible.

As to the objects, there are other build-in ones than a stroller. YOLO's default models can recognize an umbrella, suitcase, ski, snowboard, surfboard, etc, as well as a person or car. Of course, it is possible to find objects unknown to the model by fine-tuning or transfer learning as we did. Sightseeing places can count the number of whatever it would like to depending on their needs. Considering these facts, we made a practice of acquiring information with AI which was difficult to obtain from a conventional way such as using ticket sales data, or was irrelevant to the data.

7. Conclusion

In this research, in the purpose of counting the number of entering strollers, and understanding their entering patterns, which Higashiyama Zoo and Botanical Gardens were not able to do, we counted them at an entrance utilizing YOLOv5, and analyzed the collected data.

At the training and counting phase, it was revealed that it is easy to make an original model in line with the needs of the park, and that environmental arrangemetns for AI were important.

At the analyzing phase, we discovered based on data that many strollers come before noon, and that there were almost no strollers after 3 pm. This information will be helpful to determine when to invite visitors, how to obtain new ones, to grasp how many visitors are not reflected in current questionnaires, or to hold better questionnaire surveys or events for such visitors. If the park concludes that measures for strollers are effective and make the park environment more stroller-friendly, it will be their new appealing point.

For future prospects, there are two main things. First, real-time inference: Though in this research, we recorded videos and used them in our laboratory, this is not desirable in terms of personal information protection when implementing it in the real world. In order not to store personal information, processing streaming videos is required. Additionally, we are

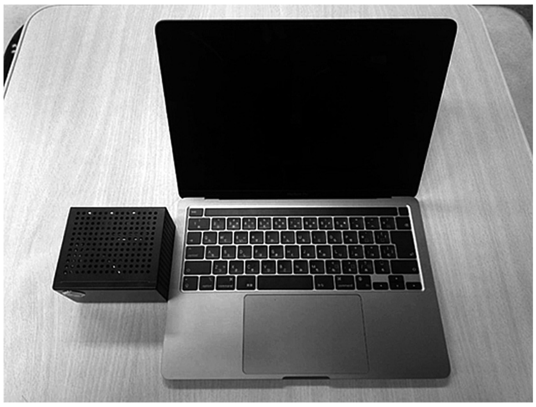

Figure 5: A smaller sized edge computer with GPU (left) and MacBook13inch (right)

considering using smaller sized edge computers (Figure 5) which are also more reasonable. This will enable a wider range of counting locations as well as the entrance we used this time. Second, counting people as well: We revealed the possibility that the strollers in the weekends and weekdays are different. We are going to make clear the average number of people passing through the gate with strollers.

Acknowledgment

We would like to thank Higashiyama Zoo and Botanical Gardens, and NEC Solution Innovators, Ltd. for their cooperation. This research was supported by JSPS KAKENHI Grant number 20K12545.

References

Arreeras, T., Arimura, M., Asada, T., and Arreeras, S. (2019). Association rule mining tourist-attractive destinations for the sustainable development of a large tourism area in Hokkaido using Wi-Fi tracking data. *Sustainability*, Vol. 11, No. 14, 3967.

Cozre (2019). Survey about travels with children (Retrieved March 31, 2021 from http://www.cozre.co.jp/blog/2403/). (in Japanses)

Iko-yo outing with kids information web sit (Retrieved March 31, 2021 from https://iko-yo.net/). (in Japanses)

Mikami, K., Kawano, M., Chen, Y., and Nakazawa, J. (2018). DeepCounter: Using deep learning to count garbage bags. *IEEE 24th International Conference on Embedded and Real-Time Computing Systems and Applications, Hakodate, Japan*, 1-10.

Morio, J., Ishi, R., Nakano, A., Hagihara, G., and Tanaka, K. (2018). Utilization of big data in urban transport planning. *IBS Annual Report Research Activities Report 2018*, 29-36. (in Japanese)

Murai, D., Hiroi, K., Yonezawa, T., and Kawaguchi, N. (2020). Analysis of behavior patterns in large-scale leisure facilities using Wi-Fi packet sensors. *IPSJ Symposium Series DICOMO Multimedia, Distributed, Cooperative, and Mobile Symposium 2020*, 766-771. (in Japanese)

Nngata, T., Morio, J., Yabe, T., Shigetaka, K., Hashimoto, H., Shibasaki, R., and Sekimoto, Y. (2016). A study about the utilization possibility of mobile spatial dynamics by the comparison with person-trip survey technique. *53rd Papers of Research Meeting on Civil Engineering Planning*, 2083-2094. (in Japanese)

Nakanishi, W., Kobayashi, H., Tsuru, T., Matsumoto, T., Tanaka, K., Suga, Y., Kamiya, D., and Fukufa, D. (2018). Understanding travel pattern of tourists from Wi-Fi probe requests. *Journal of Japan Society of Civil Engineers, Ser. D3 (Infrastructure Planning and Management)*, Vol. 74, No. 5, I_787-I_797. (in Japanese)

NTT DOCOMO (n.a.) (Retrieved April 3, 2021 from https://mobaku.jp/about/). (in Japanese)

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779-788.

Shingai, H., Imai, R., Ikeda, D., Hashimoto, T., Hayakawa Y., and Suzuki, A. (2017). Analysis on change of visitors before and after the open of high speed rail using mobile spatial statistics. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 12, 959-974.

Somusho, Skull practice with experts about EBPM secretariat (2018). Report of a skull practice about EBPM (Evidence-based policy making) (Retrieved March 31, 2021 from https://www.soumu.go.jp/main_content/000579366.pdf). (in Japanese)

Tanaka, K., Kamiya D., Fukuda D., Irohoi N., Yaginuma, H., Suga, Y., and Yamanaka R. (2019). Analysis of the tourists' travel behavior using Wi-Fi packet sensor: A case study in Okinawa Main Island. *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, Vol. 31 No. 6, 876-886. (in Japanese)

(Received April 8, 2021; accepted April 27, 2021)