Extracting Sightseeing-Related Information from Social Networking Service

SHIMUL Rakibul Hasan †, Eriko ISHII ‡, Mitsuho YAMADA ‡, Yuko HOSHINO ‡
† Department of Information and Communication Technology, Tokai University, Tokyo, Japan
‡ Kagoshima Prefectural College, Kagoshima, Japan
‡ Department of Information and Communication Technology, Tokai University, Tokyo, Japan
‡ Department of Information and Communication Technology, Tokai University, Tokyo, Japan
Email: † eng.shimulrakibulhasan@gmail.com, ‡ erikoishii@k-kentan.ac.jp, ‡ myamada@tokai.ac.jp, ‡ hoshino@tokai.ac.jp

Abstract
Technological advancement now allows potential tourists to view sightseeing locations virtually. The study focuses on building a system applying technology that can help tourists make informed decisions before traveling. First, posts on sightseeing acquire from social networking services. If a post contains an image, the objects are detected and captions. Text mining uses to identify the outline of the text if it contains the text. If a post has both images and text, then both detection and mining perform. GPS is used to obtain location information from the posts. Finally, the image captions and textual summaries combine with the location information and provide a list (automatically generated by the system) of data in a particular post about sightseeing locations.
Keywords: Objects detection, image captioning, text mining, and location detection.

1. Background

Following background, currently, one person has multiple devices such as PCs, smartphones, tablets, etc., and is browsing information on the display of various standards and sizes. In additions, research on the use of big data in various fields has increased. In Japan, many agencies have been working on the dynamic of tourism based on information from roaming data. Obtained data using GPS information and information on posts from Social networking services (SNS) in each region of Japan.

There are many ways to identify sightseeing locations and activities of interest for tourists; studying SNS posts is among them. These posts contain many images, texts, or images with textual descriptions. The numerous cultural events and cherry blossom spots attract many tourists and share their experiences on the internet using SNS.
Virtual sightseeing could be helpful in areas where multiple or lesser-known sightseeing locations and help tourists make decisions before traveling. The lack of tourist interest in these locations may be due to difficulties accessing the places or the locals not being interested.
In this paper, our aim to provide the information obtained with our method that discover lesser-known sightseeing locations. And what has changed on those locations by SNS posts.

2. Related research

In the analyses of posts to Instagram relating to sightseeing locations by Ayako Sawada. It is associated with the combination of spots and time, which visited from the place where it was made [2].

In a comprehensive Survey of Deep Learning for Image Captioning by Hossain et al. [8], the main object is extracted and discriminated from the image, and add a caption. However, it corresponds to only one object in the image.

In understanding emotions in SNS Images from the posters' perspectives by Song et al. [7], images are differentiated based on location, culture, and expressions. Different countries or even different regions in a country have their expressions, which are reflected in images.

In the information extraction from text messages using data mining techniques by Ahmad and Varma, [10], information is extracted and classified from text messages using support vector machines (SVM) and k-nearest neighbors (KNN) algorithms.

In the obscure sightseeing spots discovering system by Zhuang et al. [3], machine learning (Histogram technique) is used to discover obscure sightseeing spots by analyzing the distance view images and the close-up view images.

The real-time object detection method by Redmon et al. [6]. That combined the YOLO and R-CNN methods to detect objects from still or moving images.

The Sentiment Analysis of SNS Data using the Machine Learning Approach for the measurement of depression by Anees Ul Hassan et al. [1] is characterized by opinions using different classifiers like Positive, Negative, or Neutral by measuring the cheerfulness of a person.

3. Overview of the proposed system

Technological advancements now allow potential tourists to view sightseeing locations virtually. Object detection, image captioning, text summarization, and location detection are those kinds of technologies. Figure 1 (a) shows the scenery and congestion, etc. of a cherry blossom festival. In other words, by detecting an object in a photo, where you can see the scenery and congestion of the place the photo was taken.

Based on the result of object detection, create a caption that describes that photo. Next, get the location information from GPS. Use this location to determine the photo location. We believe that by combining the caption and location information, it is possible to extract the sightseeing location, as shown in Figure 1 (b). And use them as a recommendation for tourists.

In Japan, the cherry blossom season is a well-known festival that attracts national and international tourists. Considering this festival as an example to understand our system, both national and international tourists attend this event. They visit the cherry blossom sites and the lesser-known sightseeing spots and share their experiences by posting on the internet using SNS.
In our system, we aim to extract information using object detection, text mining, image captioning, and detect locations from sightseeing-related posts. We also examine the effects on those sightseeing locations through posts made by visitors.

(a) Detections (b) Places

Figure 1: Cherry blossom festival. In (a), Object detection, text summarization and get the location; In (b), After detections what kind of place we get from that information.

We initiate processing to extract, analyses tourism-related posts and create a post list for specific spots to calculate posts per day in each field on the internet. The procedure for the same is explained in the subsequent sections. Figure 2 shows the proposed system.

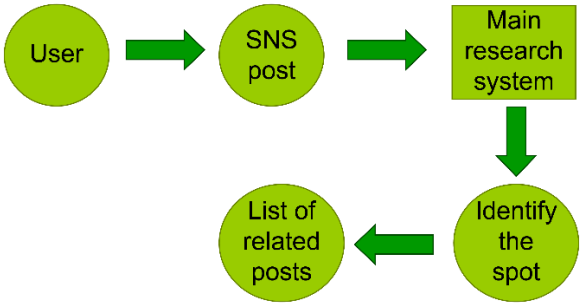

Figure 2: Proposed system diagram.

(1) To find posts on tourist information and consider the following approach. If the post contains an image, then detect objects from the image. After the detection is complete, add a caption to identify the overall description of the image. If the post contains the textual description, summarize the text from the post, using text mining to understand the text. If a certain post contains both image and textual description, we conduct both detection and text mining will perform.
(2) To analyses user's posts: first, we obtain location information of the posts based on their Google map coordinates, (longitude and latitude) as indicated by the GPS. The lacking GPS information can also be detected using gazetteers' information, which includes the geographic name and its coordinates.
(3) Finally, combine the image captioning, textual summary with location information. Moreover, provide a list (generate automatically by the system) of data in a particular post about sightseeing-related information.

3.1 Object detection

To detect objects, we would use the combined YOLO and R-CNN methods. They will even help us in real-time or in video detection [6].

3.1.1 YOLO-Object detection

This method is based on technology related to computer vision and image processing that deals with the detection of instances of semantic objects of many sectors (such as humans, buildings, or cars) in digital images and videos. Every object has some features that facilitate the identification of its —for example, all circles are round. Object class detection uses these special features. A similar approach is used for face identification where eyes, nose, and lips, skin color and distance between eyes could use.

3.1.2 Region-Based Convolutional Neural Networks (R-CNN)

R-CNN is a family of machine learning models for computer vision and object detection. The original goal of R-CNN was to take an input image and produce a set of bounding boxes as output, where each bounding box contains an object and the category (car or pedestrian) of the object. More recently, R-CNN has been extended to perform other computer vision tasks.
To instruct the system to detect objects, we will use PASCAL VOC 2007, 2012 data set, followed by the PICASO data set or People Art data set.

3.2 Text mining

Text mining is the process of information extraction from text messages using data mining techniques. Analyze the components of the posts using textual descriptions (keywords and emoticons). We can estimate a person's expression about a sightseeing location by applying data mining techniques to the posts a person shares on the internet.

To extract the information from the textual we used *support vector machines* (SVM) and *k-nearest neighbors* (KNN) algorithms for training and testing data.

3.3 Image Captioning

Many kinds of image captioning are available; we use two types of them.

One is automatic image captioning using text summarization techniques. Here is a semantic-graph-based summarizer we use for automatic image captioning.

Another is an attention-based image captioning method. The method describes the salient contents of an image automatically. They can concentrate on the salient parts of the image and generate the corresponding words similarly. This method is capable of extracting the flow of abstract meaning based on the semantic relationship between visual information and textual information. It can also obtain higher-level semantic information by proposing a scene-specific context.

3.4 Discovering Location

We are using Twitter to obtain location information which has a big collection of data. There are three layers of user timeline, which indicate the temporal information of the user location and "check-in" behavior: 1) Geographical Property of Social Connections, 2) Temporal Patterns of Geographical Check-ins, and 3) Semantic Indications of Check-in Content.
We also use Crisis response and Inferring location information in social media for better understanding the location information.

3.4.1 Crisis Response

In a crisis situation, parameters could appear one of three forms on Twitter: 1) keywords, 2) geographical bounding boxes, or 3) user name with little text, to describe the situations.
Here to identify the locations two types of system could be use, one is Tweet Tracker who enables a first responder to collect Twitter data about a crisis by specifying parameters. Another is Tweet Explorer designed to help first responders obtain situational awareness.

3.4.2 Inferring Location Information in Social Media

Only approximately 1-2% of all the tweets posted on Twitter are geotagged. Researchers have focused on uncovering the locations of users who do not share their location on social media. The location could be uncovered from three perspectives, 1) user's profile location (the location they provide in profile), 2) where they live (where the user lives now), and 3) the tweet's location (where the user tweet is from) [4]. There are two challenges in obtaining a tweet's location: 1) Language differences and 2) Geographic time differences.

4. Implementation of object detection

In recent, we are working with Object Detection and Captioning Using Tensor flow and Image AI.

The Tensor Flow object detection API is the framework for creating a deep learning network that solves object detection problems [5]. There are already pre-trained models in their framework. Which includes collections of per-trained models trained on the COCO dataset, the KITTI dataset, and the Open Images Dataset. These models can use for categories only in this dataset. Figure 3 shows the results of creating a detection program using the Tensor Flow object API and performing extraction destruction.

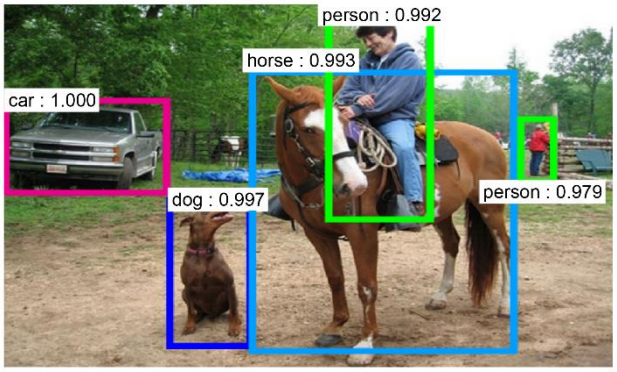

Figure 3: Detecting objects using Tensor Flow.

Image AI provides a simple and powerful approach to training custom object detection models. Using image datasets that are in Pascal VOC annotation format, also using the YOLOv3 architecture [9]. The training process generates a JSON file that maps the names of the objects in a given image dataset and the detection anchors, as well as creates many models. It detects objects in images, videos and performs an analysis. Figure 4 shows the execution result of the object detection program using Image AI.

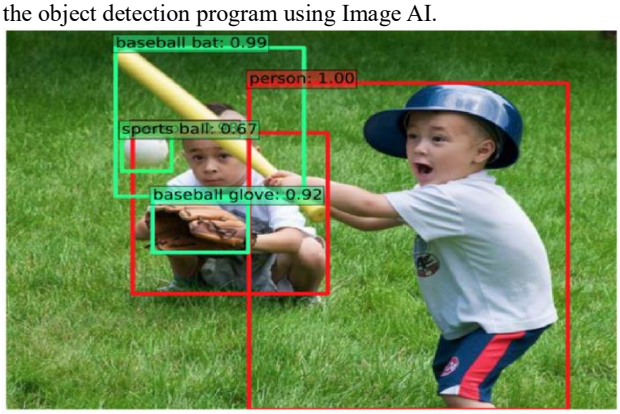

Figure 4: Detecting object using Image AI.

5. Output analysis

Advanced technology like Tensor Flow and Image AI were effective for cars and people to detecting objects [9]. However, when the objects are mixed or in a group, the result becomes poor. Even when the photo quality is not good, the accuracy of object detection becomes decreasing.

To overcome this problem, we must keep customizing them (Tensor Flow and Image AI). And increasing their accuracy for good results of any objects or photo qualities and pre-train them for detecting natural elements.

6. Future works

In the future, we will work on location information. Those are the methods such as how to detected locations from any textual descriptions, check-in data, or GPS information. And also, without GPS information using gazetteers' information to detects location. In addition, we would like to realize new attractive content by distributing related tourist information in a format that matched all kinds of devices.

7. Conclusion

This research will assist in surveying posts made by SNS data about sightseeing information. Our proposed system provides a list (generated automatically by the system) of data about tourism-related by captioning the images or summarizing the textual descriptions with location information.

Furthermore, it could use for a content-based image retrieval system. And by laying out the created tourist information according to the specifications of the user devises, the system will be easier and more understandable for the users. Therefore, people can make informed decisions about traveling while looking at the screen of their devises, share their experiences with friends, and saving time and resources.

8. Acknowledgment

Parts of this work were supported by JSPS KAKENHI Grant Numbers JP17K02129 and JP20K12411.

References

[1] Anees Ul Hassan et al, "Sentiment Analysis of Social Networking Sites (SNS) Data Using Machine Learning Approach for the Measurement of Depression" [Online], Kyung Hee University, South Korea, (2017), https://doi.org/10.1109/ICTC.2017.8190959

[2] Ayako Sawada, 'Analyses of posts to Instagram Relating to Sightseeing Locations', Hokuriku Gakuin Junior College, Journal of Global Tourism Research, and Volume 3 Issue 2 Pages 123-128, (2018), https://doi.org/10.37020/jgtr.3.2_123

[3] Chenyi Zhuang et al "Anaba: An Obscure Sightseeing Spots Discovering System" [Online], Graduate School of Informatics, Kyoto University, (2014), https://doi.ieeecomputersociety.org/10.1109/ICME.2014.6890225

[4] Fred Morstatter, Huiji Gao, and Huan Liu, "Discovering Location Information in Social Media" [Online], Arizona State University, (2015), http://www.public.asu.edu/~fmorstat/paperpdfs/ieee_deb_june_2015.pdf

[5] Hideki Shimizu, "Challenge Object Detection with Tensor Flow's Object Detection API" [Online], **(**2020**),** https://github.com/tensorflow/models

[6] Joseph Redmon et al, "You Only Look Once: Unified, Real-Time Object Detection" [Online], University of Washington, (2016), https://doi.org/10.1109/CVPR.2016.9

[7] Junho Song et al, "Understanding Emotions in SNS Images from Posters' Perspectives" [Online], Hanyang University, South Korea, (2020), https://doi.org/10.1145/3341105.3373923

[8] Md. Zakir Hossain et al, "A Comprehensive Survey of Deep Learning for Image Captioning" [Online], Murdoch University, (2018), Australia, https://dl.acm.org/doi/abs/10.1145/3295748

[9] Moses Olafenwa and John Olafenwa, "Object Detection with Image AI in Python" [Online], (2018), https://github.com/OlafenwaMoses/ImageAI/

[10] Sartaj Ahmad and Rishabh Varma, "Information Extraction from Text Messages Using Data Mining Techniques" [Online], Malaya Journal of Matematik, Volume S: 26–29, (2018), https://doi.org/10.26637/MJM0S01/05