Munenari Inoguchi, Keiko Tamura, Kousuke Uo, and Masaki Kobayashi: Validation of CyborgCrowd Implementation Possibility for Situation Awareness in Urgent Disaster Response -Case Study of International Disaster Response in 2019-, The fourth IEEE Workshop on Human-in-the-loop Methods and Human Machine Collaboration in Big data (IEEE HMData2020), Proceedings of 2020 IEEE International Conference on Big Data (Big Data), pp. 3062-3071, December 2020.

Validation of CyborgCrowd Implementation Possibility for Situation Awareness in Urgent Disaster Response -Case study of International Disaster Response in 2019-

Munenari Inoguchi
School of Sustainable Design
University of Toyama
Toyama, Japan
inoguchi@sus.u-toyama.ac.jp

Keiko Tamura
Risk Management Office
Niigata University
Niigata, Japan
tamura@gs.niigata-u.ac.jp

Kousuke Uo
School of Informatics
University of Tsukuba
Tsukuba, Japan
kousuke.uo.2019b@mlab.info

Masaki Kobayashi
School of Informatics
University of Tsukuba
Tsukuba, Japan
makky@klis.tsukuba.ac.jp

***Abstract*—At disaster response, it is essential to grab whole picture of damage situation quickly and early after disaster occurrence in order to make disaster response effective and efficient. However, it takes much time to understand damage situation because there is not enough information about it. Against this issue, we proposed implementation of CyborgCrowd for situation awareness in disaster response. In order to validate its possibility, we planned the first international disaster drill in October, 2019. In this drill, we simulated to detect flooded area by West Japan Flood occurred in 2018 from aerial photos by collaboration between crowdsourcing and AIs following Human-in-the-Loop process. Especially, in this drill, AIs were also crowdsourced. In this research, we validated the transition of the efforts from crowdsourcing and AIs to detecting flooded area, and verified the accuracy of result by comparing with the actual flooded area published by Geospatial Information Authority of Japan. Furthermore, we found some suggestion about features of detection results by humans and AIs. For example, some humans detected flooded area roughly, however AIs detected it much closely. Based on those features, we proposed the way to decrease the difference between results by humans and AIs. This was essential for local responders to understand the whole picture of damage situation after disaster occurrence urgently. In this paper, we introduced the framework of international disaster drill, clarified the result of validation, and mentioned the possibility of effective collaboration between crowdsourcing and AIs for quick situation awareness in disaster response.**

Keywords—flood, situation awareness, disaster response, CyborgCrowd

I. Introduction

Once disaster occurs, local responders has to make effective decision based on the disaster situation. In order to realize it, it is said that developing "Common Operational Picture" is necessary. However, it is difficult to gather the information in terms of damage situation as soon as after disaster. Against this issue, we proposed the concept of CyborgCrowd. CyborgCrowd is consisted of harmonious collaboration between humans and AIs to solve any kinds of problems hidden in our society. Most of those problems were caused by lack of resources, time, methodologies. In starting this project, we have no way to apply it to actual disaster response. Then, we applied it to disaster drill to design the detail of scenario for implementation firstly, and we evaluated efforts of CyborgCrowd for disaster response through some drills. In the first time, we decided to treat 2018 West Japan Flood disaster as a case study. This disaster was most severe flood disaster since the Second World War in Japan. If we solve the problem to this disaster, we can apply CyborgCrowd to other disasters.

In this paper, we introduced the general outline of the first international disaster response drill with CybrogCrowd, designed the detail of scenario for implementation, and validated the results from CyborgCrowd application. Finally, by reviewing the results with participants, we confirmed the efforts of CyborgCrowd for disaster response.

II. Related Works about Utilization AIs or Image Processing for Disaster Response

Recently AIs were spread into our society, and private companies, agencies and research institutes have tried to utilize it for effective disaster response. However, those challenge are still in the process of implementation. Furthermore, those in Japan have not been collaborated with other country by taking advantage of strength of cyber network internationally. In this section, 3 examples of those challenges in Japan are introduced.

A. Evacuation Drill with AI chatbot [1]

Council on Artificial Intelligence for Disaster Resilience implemented AI chatbot to support residents' evacuation in the disaster drill held on September 13th, 2020. In this drill, each resident can communicate with AI chatbot for getting necessary information about evacuation in LINE. This AI chatbot provide them the necessity of evacuation and information about obstacles in the way of evacuation. Furthermore, residents can register the information of damage situation by themselves to this system. This system aggregates those information, however local responders cannot detect the damaged area immediately. This challenge was limited in Japan without international collaboration.

B. *Detecting Flooded Area by AIs [2]*

Spectee Inc., which is a private company, produced "Spectee Pro". This product is a system to provide risk management information by AI in real-time. In the case study of July Rainfall disaster in 2020, Geospatial Information Authority of Japan utilized it for detecting estimated flood area. In this challenge, images posted in SNS were processed with digital elevation data. The detail of this information processing process was not published. Furthermore, this system can be used only the authorized users, thus it could be difficult for local governments to utilize it if they have not enough budget. In our project, any local governments should utilize CyborgCrowd easily, rapidly and reliably.

C. *Estimating Flooded Area by Image Data Processing*

National Research Institute for Earth Science and Disaster Resilience (NIED) can take satellite images through JAXA after disaster occurrence immediately because JAXA is a member of international disaster charter. Once large-scale disaster occurs, NIED always establish "Crisis Response Site". In the case study of Typhoon 1919 in 2019, they utilize the technology of image data processing to estimate flooded area [3], because affected area by this disaster was spread all over Japan. However, this challenge was on-going, and the result of estimation was not accurate. They have tried to improve it now.

III. Implementation of CyborgCrowd for Disaster Issue

A. *Overview of CyborgCrowd*

CyborgCrowd is proposed by Prof. Morishima [4],[5]. In the project of CyborgCrowd, they conduct research on theory and implementation of middleware that achieves flexible integration of crowdsourced processing, machine processing and any combination of them, taking the availability of people and algorithms into consideration as shown in Fig.1. They also conduct feasibility studies in various application domains including natural disaster domains, aiming to establish flexible and scalable infrastructure that aggregates human and machine intelligence.

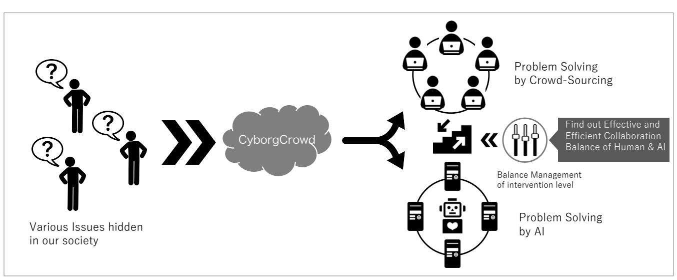

Fig. 1. Overview of CyborgCrowd

CyborgCrowd could overcome any kinds of problems lurking in our actual society. Those problems have usually the feature of "urgent", "long-tail" and so on. Against those problems it is effective for humans to process. However, it is difficult to gather a lot of humans' contribution and it takes much time. On the other hand, machines including AIs could solve the problem about the power of contribution and time-cost. However, machines are not always perfect. Therefore, it is necessary to establish the environment of collaboration between humans and machines.

In regards to disaster issue, it is urgent and it requires the accuracy of information about damage situation. However, there is still no way to utilize the humans in anywhere and AIs effectively and efficiently. We believed CyborgCrowd enable to make the current disaster response effective. Furthermore, the concept of "Society 5.0" was spread all over the world including Japan [6]. Recently, the national government has promoted to implement Society 5.0 into our actual society in Japan. Understanding this background, we can gather much more information than before, and we can train the superior AIs by using the information scattered in our society. Thus, CyborgCrowd could become effective platform for us to solve any kinds of problems in near future. In addition, there were already a lot of practical research based on the framework of CyborgCrowd [7],[8],[9],[10].

B. *Detecting Disaster Issue Using CyborgCrowd*

Disaster response is a steak of decision making (Fig. 2). In order to realize effective disaster response, it is well-known that responders have to develop "Common Operational Picture (COP)" consisted of damage situation and resource status [11]. To develop the COP, local responders should gather the information about disaster and damage situation, and aggregate it in an integrated map. However, it takes much time to gather those information, thus the disaster response tends to be late generally. The reason, why it takes much time to gather the information relating to disaster and damage situation, is that it is difficult to detect affected area comprehensively, to detect the damage status of each affected area, and to grab the number of affected people and buildings.

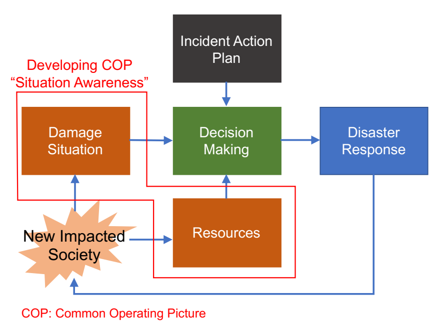

Fig. 2. Framework of decision making in disaster response

The disaster response should be done immediately and efficiently, thus COP has to be developed urgently. To overcome this issue, local responders utilized the result of simulation that was calculated before the disaster, or gathered a lot of personnel for gathering information about damage situation. However, there are some problems: 1) the result of simulation cannot fit actual damage situation generally because the damage situation rely on the scale of disaster, 2) the personnel in affected area is limited when huge disaster occurs, 3) it is difficult for personnel in other area to gather in the affected area physically at catastrophe because infrastructure could be damaged, 4) it is difficult for them to grab whole picture of damage situation when the affected area is huge such as 2011 East Japan Earthquake. Fig. 3 shows the time cost to detect building and human damage by local responders. About 22,000 persons were dead and about 400,000 buildings were damaged severely by this catastrophe [12]. At that time, it took about 1 year for them to detect almost of collapsed buildings, and about 1 month to detect almost of casualties. Anybody can understand that it was too late for them to execute urgent and efficient disaster response.

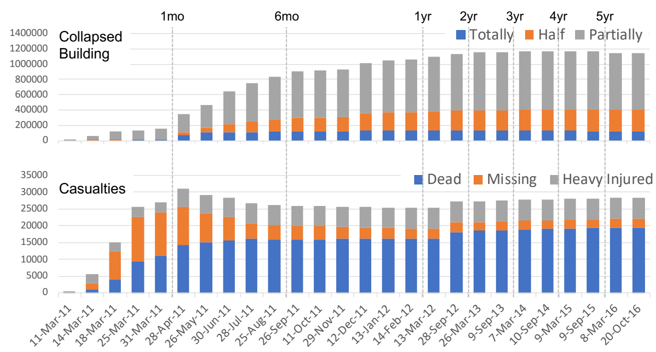

Fig. 3. Transition of detecting damage situation at 2011 East Japan EQ

The concept of CyborgCrowd is harmonious collaboration between human and AI to solve any kinds of problems. Against the issues happened in disaster response, we believed that CyborgCrowd could overcome the issue of developing COP urgently and efficiently. Especially, CyborgCrowd can solve the problem regarding to lack of human resources by utilizing crowdsourcing and the problem regarding to time cost for detecting affected area by AIs. If humans (workers) in crowdsourcing detect some affected area, AIs learn the result from crowdsourcing and AI detects other affected are by themselves, COP can be developed quickly. In the first phase of disaster response, local responders want to grab the whole picture of damage situation roughly, thus they do not require the detail of damage situation with high accuracy. Considering the status of actual disaster response, CyborgCrowd could support to develop COP in affected area. Fig. 4 shows how to implement CyborgCrowd in the phase of decision making in disaster response.

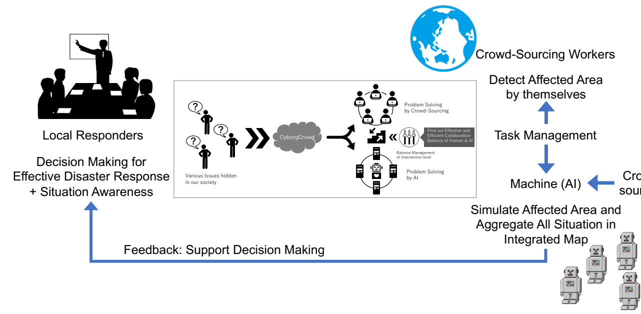

Fig. 4. How to implement CyborgCrowd to decision making at disaster

IV. Implementation of CyborgCrowd for Disaster Drill

A. *Preparation for First International Disaster Drill with CyborgCrowd*

In order to validate the possibility of CyborgCrowd implementation for disaster issue, we had started to plan the disaster drill in which local responders utilized CyborgCrowd for detecting affected area. Firstly, we determined to treat flood as a type of disaster in the drill, because we were recently attacked flood disaster for many times, and it should take much time to grab disaster situation when the flooded area is much large.

In this disaster drill, it had to be essential for international workers in crowdsourcing to involve for validating of possibility of cybrogcrowd for disaster response. Thus, we executed the process mentioned below.

1) *Dec., 2016: We proposed International Disaster Response Drill with CyborgCrowd in ICADL Panel of "Natural Disasters: What KID Can Do in/for It" in order to ask audience to participate in this project.*
2) *Jul., 2018: In order to detect what we should prepare for the disaster drill and which kinds of CyborgCrowd functions we can implement, we executed a pretest for implementation of CybrogCrowd in disaster field in Tsubame city, Niigata, Japan as "Japan Disaster Response Drill in Tsubame city".*
3) *Jul., 2018 – May., 2019: We designed scenario of international disaster drill with CyborgCrowd.*
4) *May., 2019: We visited Banda Ahce city, Indonesia in order to develop the basis for collabration in this disaster drill. Banda Ahce city was attacked by huge tsunami in 2004.*
5) *Jun-Jul, 2019: We visited Ehime prefecture in Japan in order to gain the acceptance for using their prefecture as the field of disaster drill. Ehime prefecture was affected by West Japan Flood in 2018. They experimented that they take much time to grab the damage situation at that flood disaster.*
6) *Aug., 2019: We visited Banda Ahce city, Indonesia again in order to gain the acceptance for concluding collaboration in this project.*

7) *Sep., 2019: We held international workshop of "Human-in-the-loop Big Data and AI: Connecting Theories and Practices for a Better Future of Work" supported by National Institute of Informatics in Japan, and explain our project in order to increase workers in crwodsourcing all over the world.*
8) *Sep., 2019: We visited Ehime prefecture again in order to finalize the detail of disaster drill scenario.*
9) *Oct. 8, 2019: We execute the first international disaster drill in Ehime connecting all over the world through CyborgCrowd. We will describe the detail of this drill in the section C.*

B. Design of the Disaster Drill Scenario

In this project, we decided that 2018 West Japan Flood disaster was the subject of a case study. 2018 West Japan Flood disaster was caused long-term rainfall in July, 2018, which was occurred by a seasonal rain front. By this disaster, 237 people were dead, 8 people were missing and 309 people were injured heavily. In terms of building damage, 6,767 housings were totally collapsed and 11,243 housings were half collapsed. In addition, 7,173 housings were affected by inundation above floor level and 21,296 housings were affected by inundation below floor level [13]. This disaster was most severe one since the Second World War in Japan. Furthermore, the flood occurred in various regions concurrently. Because of these circumstances, it was too difficult for responders to detect the affected area and to understand damage situation rapidly. Furthermore, in this case study, we selected Mabi district in Kurashiki city in Okayama prefecture where was totally inundated at the disaster. The participants as local responders were staff of Ehime prefecture, and they knew well already the damage situation in Ehime prefecture, therefore we selected Mabi district.

The objective of this disaster drill was how to collaborate between humans in crowdsourcing and AIs in cyber area in order to detect flooded area. Then, for requesting humans' tasks, we utilized Crowd4U as the platform of crowdsourcing. this platform had a lot of past performance, and it can be used anyone under academic objectives. However, it is not ensured that workers in Crowd4U are prepared for the execution of all tasks immediately. Therefore, we decided to utilize Yahoo! JAPAN Crowdsourcing and Amazon Mechanical Turks additionally. In terms of tasks for workers, we prepared 4 selections of answers: 1) non-flooded, 2) all fully flooded, 3) partially flooded, and 4) covered with clouds. In previous research, we found that some aerial photos included the clouds, so we add "covered with clouds" as an item of answer. Fig. 5 shows the description of a task implemented in crowdsourcing in this case study.

In the next place, we had to prepare the materials for them to detect flooded area. As materials, we used aerial photos corrected to vertical images that were 106 images. Those photos covered over Mabi district, and were published by Geospatial Information Authority of Japan after the disaster. Using those images, we prepared the tasks to classify images to

Fig. 5. A task implemented in crowdsourcing to detect flooded area

"flooded" or "non-flooded" in crowdsourcing. In previous research, we tried to detect damaged buildings from aerial photos by crowdsourcing, however the affected area was too small for workers to get the answer of "damaged". Before detecting affected area, there is no way to detect which images include affected area. In actual, those vertical images covered over the not-flooded area as shown in Fig. 6. The flooded area detected in official was only 3.5% to the area covered with 106 images.

Against this issue, we proposed the processing flow for task assignment as described below.

1) Assign tasks with the aerial photos with original size.
2) If the answer of task is "all flooded" or "non-flooded", then the answer was treated as final answer.
3) If the answer of task is "partially flooded", then the image is divided to 4 small segementation (2 by 2) automatically.
4) Processes of #2 and #3 are continued until the answers are "all flooded" or "non-flooded".

These processes are shown in Fig. 7. Especially, understanding the size of segmented images which can including a common building, the division process is repeated more than 5 times in this case study.

On the other hand, we utilize AIs. The AIs are also crowdsourced. It is usual that we define the specification of AI when we utilize or develop AI. However, considering the circumstances of disaster, we cannot design the spec of AI before because the features of disasters are different each other. When AIs gather from crowdsourcing to our project, each AI starts to be trained with the teacher data consisted of human workers' answers by itself. Then, AIs are trained by itself with those data continuously, and AIs detect flooded area periodically.

In this time, CyborgCrowd has 2 types of resutls: one is results from humans in crowdsourcing, and the other one is results from AIs in cyber environment. If CyborgCrowd shows 2 types of results to local responders in parallel, it is difficult for them to understand damage situation by integrating those results by themselves. Against this issue, we have to decide how to integrate them in a map as COP. Since this case study is first attempt, we visualize those results to staff of city and prefecture as local responders each other in a screen, and we will discuss how to integrate is most effective for local responders in developing COP through reviewing and validating process.

Fig. 6. Comparison between coverage of actual flooded area and photographing range

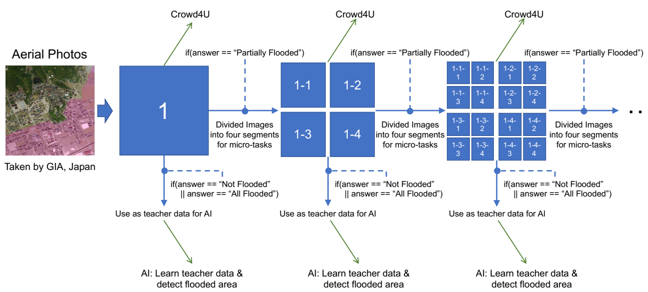

Fig. 7. Process of image division for tasks in crowdsourcing

C. Exacution of the Disaster Drill

On October 8th, 2019, main members of us gathered in Ehime prefecture, and we connected all over the world including the other area in Japan through CyborgCrowd. We started the disaster drill at 10:00 a.m. Firstly, we set aerial photo over the affected area by 2018 West Japan Flood disaster in CyborgCrowd. CyborgCrowd started to divide the images as tasks for crowdsourcing, then we asked tasks through Yahoo! Japan crowdsourcing and Amazon Mechanical Turk in the world. In Ehime prefecture, local responders (city staffs) were watching the detection result of flooded area on the screen updated in real time that each segmented image is the image of flooded area or one of non-flooded area. Especially, in Banda Aceh city, Indonesia, about 10 students supported this project and they executed the tasks in crowdsourcing (Fig. 8). Parallelly, AI learned from humans answer in crowdsourcing, and classified all segmented images to flooded or non-flooded. Then, CyborgCrowd periodically aggregate both of results by human and AI into the map of damage situation. At noon, the AI brought out the result of flooded area detection, then all staff gain the first damage situation map. After that, they monitored the transition of clarification of damage situation as a map.

In the afternoon on that day, they keep monitoring the process of clarification of damage situation by human and AI. At 15:00, the transition of damage situation was almost stabled. Then, we described the technical methodology of CyborgCrowd for detecting damage situation at disaster (fig.9), and we reviewed the result of flooded area detection with local responders in Ehime. Finally, all of us, including staff from headquarter in Japan, Indonesia and Philippine, reviewed the implementation of CyborgCrowd project for damage situation awareness in urgent disaster response. In this review, CyborgCrowd for damage situation awareness was received a high evaluation generally. This review was qualitative. In the next chapter, the accuracy of results was validated quantitively.

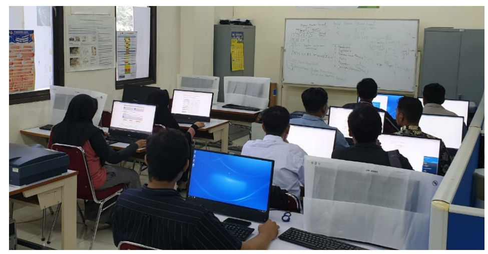

Fig. 8. Participants as workers in crowdsourcing at Syiah Kuala University in Banda Aceh, Indonesia

Fig. 9. Participants as responders in virtual disaster operation center at Ehime prefectural office, Japan

V. Validating the Result of CyborgCrowd Implementation

Through this disaster drill, we recorded the activity logs of the classification for each segmented image by human and AI in chronological order. In this chapter, we will validate the result of flooded area detection in five viewpoints: 1) how much humans were contributed to detection of flooded area, 2) how reliable results by humans were for detection of flooded area, 3) how reliable results by AIs were for detection of flooded area, 4) how possible integrated results can detect flooded area, and 5) how CyborgCrowd can be useful for responders and supporters in disaster response.

A. *Preparation for Results Validation*

We gathered log data from cybrogcrowd platform in which the task logs by human in crowd4U as crowdsourcing and the classification logs by AIs were included. As mentioned in chapter III, number of the smallest segmented images used in tasks in crowdsourcing was 108,544. We selected 106 aerial images covering over the flooded area and the non-flooded area. Each image was divided to 32 by 32 segmented images as smallest, thus the total number was 108,544. AI classified each of all those images to 4 classes that were "flooded", "non-flooded", "partially flooded" and "covered with clouds".

The result of classification by AI was managed with index of each image in each time-phase. However, human started to judge images with those original size firstly. Only when the answer was "partially flooded", the image was divided into 2 by 2 segmented images, and the segmented images were set as tasks in crowdsourcing again. Thus, if human judged some images as "flooded", the relevant images were not segmented and the result was managed with the not-segmented images comprehensively. This means that the sizes of answer image by human and AI were different. In order to compare between results by human and AI, we had to align the size of answer images to smallest ones. Then, we distributed the answers by humans in not-segmented images to smallest segmented images, and we unified the number of answers by human and AI to 108,544.

When we validate the results of CyborgCrowd, we had to prepare the correct answer. In this disaster drill, we tried to detect flooded area by 2018 West Japan Flood, however nobody knows the actual flooded area in detail. Then, we decided to use the flooded area detected by the Geospatial Information Authority of Japan (GIA) [14] as correct answer. In this chapter, we compared the flooded area detected by CyborgCrowd and GIA in order to validate the accuracy and reliability of the results by CyborgCrowd.

B. *Validation-1: how much humans were contributed to detection of flooded area*

After gathering the result of detection by humans as raw data, the size of image of each result was not unified. Then, if we found the result of image with big size, which means that the image was classified before being divided to smallest segmented image, we put the same answer to the smallest segmented images covering over the image which humans answer to. After we distributed the answers to big size images to smallest images, we surveyed that each smallest image were placed in flooded area or not by comparing with the area detected by GIA. Finally, we gained the result of validation at every hour shown in Fig. 10.

At the beginning of disaster drill, no worker performs any task because CyborgCrowd did not assign any task yet. After that, CyborgCrowd started the assignment of task and workers has participated in the project. After 1 hour from the beginning of disaster drill (at 11:00), only 9.4% of smallest segmented image sets were processed. This means only 9.4% of the area was judged as "flooded" or "non-flooded". However, in those results, the results in the class of "partially flooded" were included. The images with those result divided to 2 by 2 segmented images for next tasks in crowdsourcing. In the early phase, the images were not segmented, thus the processing rate increased immediately when a few images were judged. Furthermore, images in other area were assigned to workers concurrently. After 2 hours from the beginning of disaster drill (at 12:00), only 73.5% of the area was processed. As this, the occupancy rate of processed images represents the degree of humans' contribution. After 12:00, there were no big change in the degree of humans' tasks contribution.

As this result, only in the early phase (about 2 hours), the contribution increased rapidly, however the contribution had not increased already after 2 hours. When the result of detection of flooded area has high accuracy, we can evaluate that the contribution by humans was enough in this project. If so, we can understand that the AI was trained enough and it could work more than humans.

Especially, in this validation, we clarified the spatial contribution by humans shown in Fig. 11. Fortunately, actual flooded area was included in the assigned area as tasks to workers in this disaster drill. In addition, humans tend to judge the images roughly. Therefore, images covered over almost of actual affected area were processed in the early phase. This caused that the training data was prepared for AI adequately.

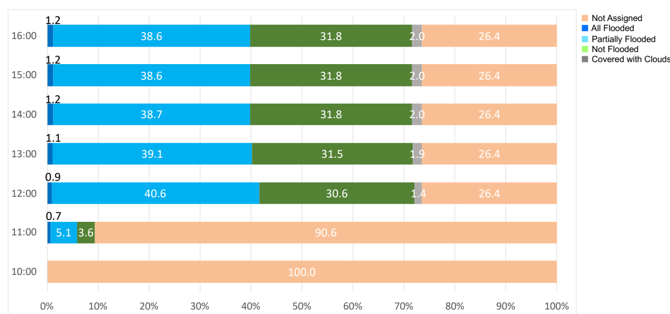

Fig. 10. Validation of humans' contribution for detection of flooded area

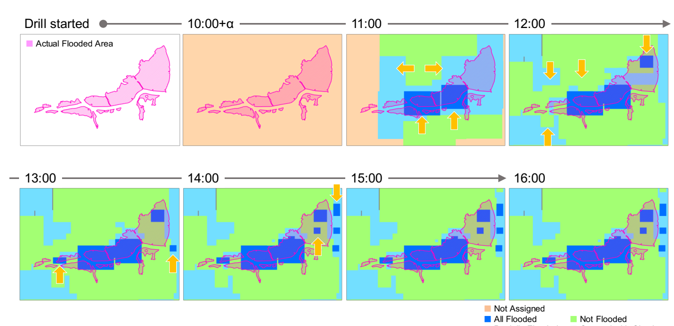

Fig. 11. Validation from the view point of spatially analysis

C. *Validation-2: how reliable results by humans were for detection of flooded area*

In the next place, we validated the reliability of results of judgement by humans in crowdsourcing. In this validation, the results were classified to "non-flooded area" and "flooded area” in correct answer data created by GIA. Fig. 12 shows the result of validation.

In the early phase, humans processed images almost in flooded area, and their answers were also correct generally. After 2 hours from the start of disaster drill, 99.5% of images taken the actual flooded area were assigned, and about 85% of answers were “flooded” or “partially flooded”. On the other hand, 72.4% of images taken the actual non-flooded area were assigned, and about 57% of answers were “non-flooded”, “partially flooded” or “covered with clouds”. If images classified to “partially flooded” were considered as correct answers in both of actual flooded area, the accuracy rate of humans’ answer for “flooded area” was very high. However, the accuracy rate of the answer for “non-flooded area” was just 57% even if it included the answer of “partially flooded”.

We consider the reason why it happened was that human tends judge roughly. Thus, even if an image includes a few of non-flooded area, human tends to judge as not “non-flooded”, but “flooded”. Because of this, some area at a constant rate were detected as incorrect answer in validating in the scale of smallest segmentation even if they intended to answer correctly.

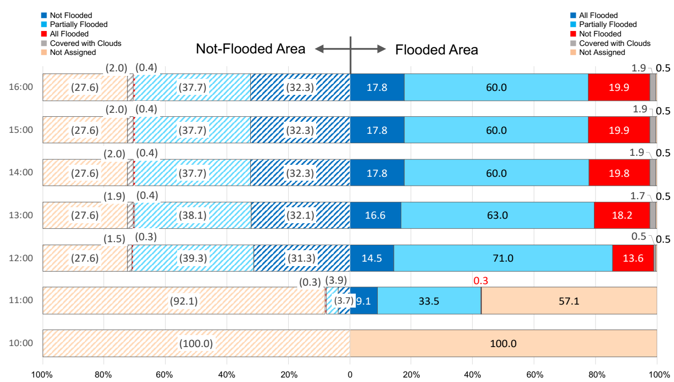

Fig. 12. Validation of accuracy of humans’ answers in actual flooded/not-flooded area

D. Validation-3: how reliable results by AIs were for detection of flooded area

We validated the reliability of results of judgement by AIs in CyborgCrowd in same ways as validation of humans’ results. Especially, AIs classified all of smallest segmented images in three times during disaster drill. In this validation, we focused on those three phases. Fig. 13 shows the result of validation.

In first detection, only 27.9% of flooded area was detected correctly while 98.7% of non-flooded area was detected correctly. We presumed the AIs were not trained enough because there were only a few training data from crowdsourcing at the beginning of disaster drill. In this drill, it took about 2 hours for AIs to classify all of smallest segmented images to 3 classes of “flooded”, “non-flooded” and “covered with clouds”. Furthermore, we clarified the transition of timeline of training from humans’ answers and classifying images into 4 classes, and visualized the results spatially as shown in Fig. 14.

In second detection, 32.4% of flooded area was detected correctly. The ratio increased a little more than one in first detection. We presumed the AIs were trained a little more because more training data was prepared by crowdsourcing. In non-flooded area, 98.6% of the area was detected correctly. The ratio does not almost change from one in first detection.

In third detection, 44.2% of flooded area was detected correctly. The ratio increased about 10% compared with second detection. Furthermore, the AIs classified some images to “covered with clouds”. This was presumed that the AIs learned the features of images in the class of “covered with clouds”. In non-flooded area, 90.5% of the area was detected correctly. The ratio decreased about 8% compared with second detection. However, 7.8% of the area was detected as “covered with clouds”. Thus, the total ratio of “non-flooded” and “covered with cloud” does not almost change from one in first and second detection.

Through this validation of results, AIs has been trained to a certain degree for non-flooded area and the area covered with clouds. However, by the transition of ratio of correct answers, the knowledge of AIs was still at a developmental stage. In this case study, we were able to ensure 5 hours for disaster drill including lunch time. If humans prepare the correct answer data more as training data, the AIs could be developed more intelligently. As mentioned in chapter II, the COP could be developed by using both of the result of humans and AIs in CyborgCrowd. In next section, we validated the possibility of utilization of CyborgCrowd for developing COP by integrating all results of detection from humans and AIs.

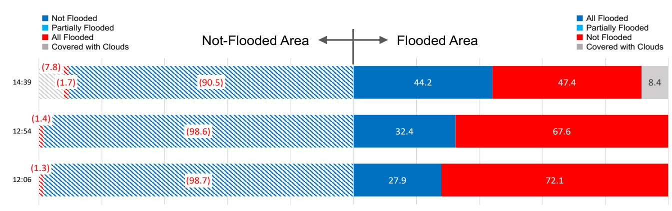

Fig. 13. Validation of accuracy of AIs’ detection in flooded/not-flooded area

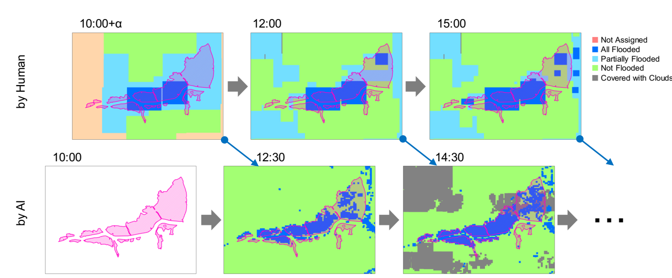

Fig. 14. Relation between humans’ results and AIs’ results

E. Validation-4: how possible integrated results can detect flooded area

In previous 3 section, we validate the results from humans and AIs each other. While the results from humans had high accuracy, it was necessary for workers to gather for tasks and to be asked continuous participation. While it took a short while for AIs to find flooded area out at once, the accuracy of results was not so high. Against this issue, we had to consider how the results from humans and AIs should be integrated effectively for developing urgent COP in disaster response.

Firstly, we categorized the combination of answer class by humans and AIs as shown in Table 1. When the class of answers from humans and AIs was matched each other, we can consider the class of processed area could be correct. If the class was different each other, the result of judgement should be controversial. In this project, we consider the judgement by humans is probably more accurate than one by AIs. Based on this consideration, we set the color of symbols in a map following the based color with flooded status.

TABLE I. INTEGRATION PATTERN WITH RESULTS OF HUMANS AND AIS

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="4">Answer by **Human**</td></tr>
<tr><td>All Flooded</td><td>Partially Flooded</td><td>Not Flooded</td><td>Covered with Clouds</td></tr>
<tr><td rowspan="3">Answer by **AI**</td><td>All Flooded</td><td colspan="2">Reliable Judge "Flooded"</td><td>Controversial Judge Probably "not flooded"</td><td rowspan="2">Controversial Judge Probably "with clouds"</td></tr>
<tr><td>Not Flooded</td><td rowspan="2">Controversial Judge Probably "flooded"</td><td>Controversial Judge Probably "not flooded"</td><td>Reliable Judge "Not Flooded"</td></tr>
<tr><td>Covered with Clouds</td><td>Controversial Judge Probably "with clouds"</td><td>Controversial Judge Probably "not flooded"</td><td>Reliable Judge "Unknown"</td></tr>
</table>

After setting the color of symbols, we reproduced a COP map representing the flooded area by integrating results from humans and AIs. Fig. 15 shows the transition of detecting flooded area as a COP.

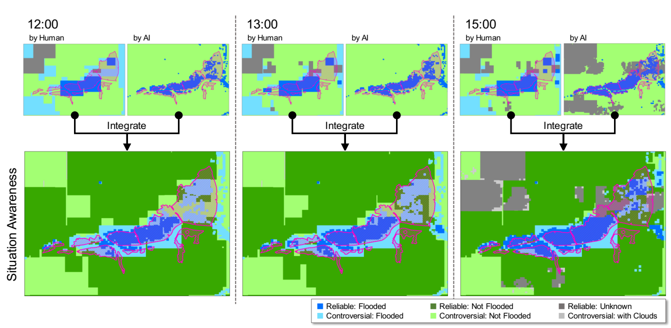

Fig. 15. Result of integration of humans' results and AIs' results in a map

At 12:00 when AIs in CyborgCrowd produced the first result, there were some area with controversial judgement of "non-flooded" in actual flooded area, however flooded area including controversial results covered almost over the actual flooded area. At 13:00 when AIs produced the second result, some area with controversial judgement of "non-flooded" in actual flooded area changed to "non-flooded" without controversial judgement. We presumed that there two primary reasons of this change: 1) humans judged in error or humans judged "non-flooded" roughly although the images included some flooded area, 2) humans judged some area as the status of "covered with clouds" and the judgement between humans and AIs was different. At 15:00 when AIs produced the third result, almost of the area with controversial judgement in actual flooded area was removed. The reason of this was that AIs classified the area with controversial judgement to "covered with clouds", and the result of humans as "covered with clouds" corresponded to one of AIs. Furthermore, the area detected as flooded was refined in actual flooded area, and almost outline of the flooded area detected by CyborgCrowd was fit to the outline of actual flooded area. At this time, almost area in actual flooded area was detected as "flooded" or "covered with clouds", and local responders could grab the whole picture of flooded area.

Finally, it took about 5 hours for CyborgCrowd to detect flooded area, and the accuracy and resolution has a certain degree. Based on this validation, we concluded that CyborgCrowd can satisfy the requirement of developing COP in urgent phase of disaster response. This was not in the final stage, thus we have to improve the accuracy of classification of segmented images much more from now. Furthermore, we have to discuss with responders who had the experience of actual disaster response in order to improve the effort of CyborgCrowd for the issues in disaster management.

F. Validation-5: how CyborgCrowd can be useful for responders and supporters in disaster response

In order to understand the usability of CyborgCrowd in disaster response, we reviewed the disaster drill with local responders and support staff in Indonesia and Philippine. This was based on a qualitative evaluation. The comments gained in the review were described below. Especially, circle as a mark of line represents positive comment and cross one represents an issue that CyborgCrowd has to solve.

1) *From officers of Ehime prefecture as local responders*

- O In order to detect which cities and towns under our prefecture should be supported, we must grab the whole picture of damage situation quickly in the first step.
- O We expect to utilize this result to estimate damage situation in order to demand support to other local governments and national government.
- O We can understand how we are affected by comparing with other prefectures. This means we can appeal our severity by disaster damage to national government to gain more support.

2) *From officers of Banda Aceh city as support staff from outside of affected area*

- O CyborgCrowd system is good system that can be implemented in disaster management.
- O The appropriate or good system for example CyborgCrowd system can be used to strengthen the decision-making during emergency response.
- O The idea of CyborgCrowd is very useful.
- × However, human and AI should work parallel. We should take care about the rapid development of technology in disaster management without considering

on social and cultural aspect. Social cultural aspect will also play important role in preparedness and mitigation cycle in disaster management system.

Generally, we gained the positive comments to the implementation of CyborgCrowd into disaster response from all members participating to this disaster drill. However, some issues were addressed. In regards to social cultural aspect, we had believed that CyborgCrowd could be utilized beyond it, however it was indicated as an issue to be solved. Furthermore, some staff indicated the necessity of education for utilizing CyborgCrowd. In near future, we should find some ways to overcome those issues out, and develop some documentations and education system to let any local responders understand the meaning of results from CyborgCrowd easily and certainly.

VI. Coclusion

Once disaster occurs, effective and efficient disaster response is required. From the aspect of social science, disaster response is a series of decision making. In order to make rational decision, it is necessary for local responders to develop "Common Operational Picture (COP)" including damage situation and the status of resource availability. However, it is difficult for them to develop the COP rapidly because of lack of human resources, information about damage situation and enough time. On the other hand, prof. Morishima proposed "CyborgCrowd" which can solve any kinds of problems hidden in our society by harmonious collaboration between humans and AIs. Just now, CyborgCrowd is developing for application to actual field. In this research, we tried to apply it to disaster response, and explored the possibility of CyborgCrowd implementation to disaster issues.

In this paper, we introduced the overview of CyborgCrowd framework, and disaster issues which can be solved by CyborgCrowd. Then, we design the workflow for implementation it to disaster issues. We implemented it to the first international disaster response drill under the participation from supporters all over the world. In this drill, we tried to detect flooded area by CyborgCrowd rapidly in a case study of 2018 West Japan Flood, which was most severe disaster since the Second World War in Japan. After the disaster drill, we reviewed this project with staff of city and prefecture and other participants, and we validated the accuracy of all results from humans and AIs. By this validation, it was found that it took about 5 hours for CyborgCrowd to detect flooded area, and the accuracy and resolution has a certain degree. Finally, we examined how to integrate those results effectively, and concluded our research.

While we mentioned that it took about 5 hours for CyborgCrowd to detect flooded area, this time-cost was required after the vertical aerial photos as materials were published. We have no information about when GIA had published those images at 2018 West Japan Flood disaster. In regard to Typhoon #1915, GIA has not published the vertical aerial photos yet now, although Typhoon #1915 caused long-term power outage and a lot of severe damage to housings. This means that implementation of CyborgCrowd in this paper relies on the timing of images publication by GIA. In Japan, the occurrence of Nankai megathrust earthquake and Tokyo metropolitan earthquake is presumed in Japan. If these earthquakes occur, it is anticipated that it takes much time for GIA to publish vertical aerial photos covering over all affected area, and CyborgCrowd cannot be started immediately.

After the case study mentioned in this research, we have started to utilize other kinds of images instead of those images published by GIA, such as images posted into SNS. However, all of the images posted into SNS are not always related to damage situation and all of them do not have geolocational information entirely. In utilizing those images, we have to solve those issues. Furthermore, the workflow developed in this paper should be modified to be fit to new kinds of images. We are planning to solve those issues, and we are eager to overcome the anticipated catastrophe in near future.

Acknowledgment

The authors would like to thank the all staff in Ehime prefectural office for cooperating with our proposed system for effective roof damage inspection although they were affected by severe flood actually, and Syiah Kuala University and Banda Aceh city from Indonesia, University of Philippines from Philippine, Yahoo! Japan Crowdsourcing, National Research Institute for Earth Science and Disaster Resilience, and Geospatial Information Authority from Japan for collaboration for this project. Furthermore, the authors would like to thank the all staff participating to CyborgCrowd project in the world. This study was conducted with support from "1. JST, CREST (#JPMJCR16E3)" and "2. Contracted research with the Ministry of Education, Culture, Sports, Science and Technology: The Tokyo Metropolitan Resilience Project".

References

[1] Counsil on Artificial Intelligence for Disaster Resilience, "Promote Indivdiaul/Mutual Aid among residents in apartments at disaster by SNS", https://caidr.jp/data/2020-09-11press.pdf, 2020.

[2] Spectee Inc., "Utilization of Spectee Pro for developing flood estimation map produced by Geospatial Information Authority of Japan", https://spectee.co.jp/200715/, 2020.

[3] National Research Institute for Earth Science and Disaster Resilience, "Crisis Response Site", http://crs.bosai.go.jp/ (accessed on November 9, 2020).

[4] Atsuyuki Morishima, et.al., "JST CREST CyborgCrowd Project", Proceedings of 1st Project Networking Workshop collocated with 6th AAAI Conference on Human Computation & Crowdsourcing, pp.2, 2018.

[5] Morishima Atsuyuki, Inoguchi Munenari, Tajima Keishi, Kitahara Itaru, "Digital Knowledge Era and the Future of Work: Challenges and the Current Status of JST CyborgCrowd Project", Workshop on Addressing SDGs Through Data Analytics and AI, International Conference on Digital Landscape 2019, pp.68-78, 2019.

[6] Cabinet Office of Japan, "What is Society 5.0" (in Japanese), https://www8.cao.go.jp/cstp/society5_0/ (accessed on September 30, 2020).

[7] Flavia Fulco, Munenari Inoguchi, Tomoya Mikami, "Cyber-Physical Disaster Drill: Preliminary Results and Social Challenges of the First Attempts to Unify Human, ICT and AI in Disaster Response", 2018 IEEE International Conference on Big Data (Big Data), pp. 3495 - 3497, 2018.

[8] Shishido Hidehiko; Kitahara Itaru, "Calibration of Multiple Sparsely Distributed Cameras Using a Mobile Camera", Journal of Sports Engineering and Technology, 234(1), pp.37-48, 2019.

[9] Rumi Hiraga, et.al., “First Evaluation of Information Support of everyone by everyone for everyone TimeLine (ISeeeTL) applied to Deaf and Hard of Hearing People Watching Sport”, The 15th International Conference of the Association for the Advancement of Assistive Technology in Europe (AAATE 2019), pp.S135-S136, 2019.

[10] Masaki Kobayashi, Kei Wakabayashi, and Atsuyuki Morishima. “ Quality-aware Dynamic Task Assignment in Human+AI Crowd”. In Companion Proceedings of the Web Conference 2020, pp.2, 2020.

[11] Haruo Hayashi and Munenari Inoguchi, “What Should Be Considered to Realize ICT Support for Effective Disaster Response and Recovery?”, IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E98-A, No.08, pp.1594-1601, 2015.8

[12] Cabinet Office of Japan, “About 2011 East Japan Earthquake (as to March 10, 2020)”, http://www.bousai.go.jp/2011daishinsai/pdf/torimatome20200310.pdf, pp.154, 2020.

[13] Cabinet Office of Japan, “About Damage Situation by Heavy Rainfall Disaster in July, 2018 (as to January 9, 2019)”, http://www.bousai.go.jp/updates/h30typhoon7/pdf/310109_1700_h30typhoon7_01.pdf, pp. 204, 2019.

[14] Geospatioal Information Authority of Japan, “Information regading to Heavy Rainfall Disaster in July, 2018”, https://www.gsi.go.jp/BOUSAI/H30.taihuu7gou.html, 2018.