Research on Dynamic Attention State during Cognitive Rehabilitation with Cooking for Patients with Acquired Brain Injury

Sho Ooi, Kazuki Hashimoto, Haruo Noma, and Mutsuo Sano

***Abstract*—Patients with acquired brain injuries present adverse symptoms such as attention, memory, and functional disorders. They prevent them from effectively executing the activities of daily living. In the rehabilitation of patients with acquired brain injuries, patients need to be aware of their cognitive states. Its effective method is to watch the experience videos and present quantitative cognitive status to patients. Current methods to evaluate cognitive states using require special/specific toolkits, and the methods are rigorous when applied for real-time dynamic evaluation. Moreover, patients are often burdened by the need to undergo tests as required by the evaluation methods. In this paper, we propose a method to evaluate the attention function from the state of handling a kitchen knife that includes dangerous movements even during cooking. As a result, we defined four attention levels during cutting behavior in cooking and could classify an average accuracy of 74.3 percent.**

***Index Terms*—Attention state estimation, cognitive rehabilitation with cooking, human support system, machine learning.**

I. Introduction

Patients with acquired brain injuries present adverse symptoms such as attention, memory, and functional disorders, as well as aphasia, which prevents them from effectively executing activities of daily living [1], [2]. In Japan, approximately half a million people live with brain injuries, and this number is increasing every year [3]. Rehabilitation against their cognitive function is called cognitive rehabilitation, there are various methods depending on the cognitive function. One of the effective rehabilitation contents is cognitive rehabilitation with cooking [4]. As a reason for that cooking includes various cognitive functions such as thinking about a process, focusing on the works (e.g. using fire and using cutlery), doing the parallel works.

There is one other thinking that is important for cognitive rehabilitation. It is that to make the cognitive training of the patients smoother in cognitive rehabilitation, it is necessary for the patients to be conscious of their own diseases [4]. This is most important for awareness and is indicated by the neuropsychological pyramid in Rusk's hospital program [5]. In order to realize their thinking, we have proposed a series of rehabilitation cycles, and have verified the effectiveness of our system as shown in Fig. 1 [4].

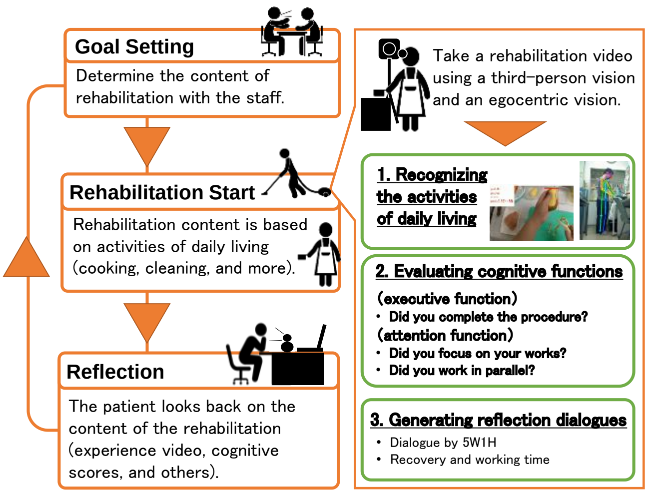


Fig. 1. Magnetization as a function of applied field.

Here, in order to aware conscious of their own diseases, it is necessary to reflect on the experience video and quantitatively score of cognitive state for the patient but it is difficult real-time to evaluate using the conventional tests such as clinical assessment for attention (CAT) [6], clinical assessment spontaneity (CAS) [6] and trail making test (TMT) [7]. Those tests are that patients and helpers are burdens because those tests are paper type test or test with a special tool, and will need time to take the test.

This paper focuses a cognitive rehabilitation with cooking that is effective in improving cognitive function, patients and helper are to decrease burdens, and this study aims dynamic to grasp an attention state during cooking. Specifically, this study focuses on cutting behavior with a kitchen knife, recognize hands state and cooking situation using YOLO algorithm [8] and Open Pose [9] algorithm, which estimates the attention state from the hands state and cooking situation.

II. Related Works

Ooi et al. focused on evaluating sustained attention during cooking behavior, suggested the difference of behavior between giving and not giving cognitive load tasks on students [10]. In their study, the group with the cognitive load had more moving the hand without the kitchen knife and opening hands than the other group. In other words, the group with the cognitive load could not focus on the cooking task.

Furthermore, they focused on evaluating distributed attention during cooking behavior, proposed a ”focus map” that integrated the brain processing model with the

Manuscript received March 9, 2021; revised July 22, 2021. This work was supported by JSPS KAKENHI Grant Numbers JP 19K20750.

Sho Ooi, Kazuki Hashimoto, and Haruo Noma are with Ritsumeikan University, Shiga, Japan (e-mail: SHO.OOI@outlook.jp, khashimoto@mxdlab.net, hanoma@fc.ritsumei.ac.jp).
Mutsuo Sano is with Osaka Institute of Technology, Osaka, Japan (e-mail: Mutsuo.sano@oit.ac.jp).

doi: 10.18178/ijmlc.2022.12.3.1087

conventional saliency map [11]. To verify a model, they conducted a multitasking cooking task under the same conditions as in reference 10; the group with the cognitive load could not focus on the cooking task than the other group. In other words, the group with the cognitive load could not focus on multiple cooking tasks.

Ji *et al*. made a virtual reality game to assess attention function [12]. In this game, they use the features as eye movement data and operation information, analyzed using the gaussian mixture model (GMM) and dynamic time wrapping (DTW), and suggested the possibility of analyzing sustained attention.

III. Evaluation Method of Dynamic Attention State

A. About the Attention Functions

Attention is classified into bottom-up attention (passive attention) and top-down attention (active attention) in neurophysiology. Bottom-up attention is unconscious attention, such as reaction to sound, reaction to moving objects, and so on. On the other hand, top-down attention is conscious attention [13], [14]. Furthermore, various researchers have studied the classification of attention function [15], [16], Kashima et al. have classified attention into the four items of i) strength, persistence, range, ii) selectivity, concentration, stability, iii) convertibility and mobility, and iv) controllability and divide [15]. Based on the four attention, Table I shows cases of attention states when cooking.

TABLE I: Kinds of Attention States When Cooking

Kinds of attention	Instruction and cases
Controlled attention	If the water boils while cutting the food ingredients, stop the cutting work, and turn off the heat.
Selective attention	Select the utilizing cookware from multiple cookware.
Sustained attention	Focus on work until the end.
Distributed attention	Do multiple tasks in parallel.

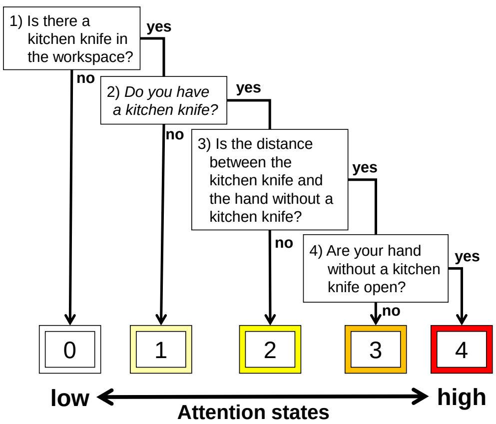


Fig. 2. The process to detect attention states.

B. Classification of Attention States during Cutting Behavior

We discussed with professors and doctors at Keio University and classified the attention states as shown in Fig. 2. In particular, this study focuses on sustained attention during cooking with cutting behavior. In addition, this study sets and fixes the camera to the top of the kitchen.

1) Is there a kitchen knife in the workspace?

In order to check to exist a kitchen knife in the workspace, it is necessary to object detection. Researchers have proposed several methods for detecting objects [8], [17]-[19], but this study uses the YOLO algorithm that provides high speed and high accuracy [8]. Fig. 3 shows a result using the YOLO algorithm during a cutting scene. This result is detected as a kitchen knife by the YOLO algorithm and is draw a bounding box as a detected object region on the image.

Fig. 3. Part of scene a detected object by YOLO algorithm.

2) Do you have a kitchen knife?

The operation of handling a kitchen knife involves dangerous situations such as injury; therefore, a system needs to make sure that the user is using the knife. In order to check to use a kitchen knife by a user, this study focuses on a feature of a kitchen knife movement.

The center of gravity of bounding box $G_t(x_t, y_t)$ when a detected a kitchen knife at current time t, the center of gravity of bounding box $G_{t-1}(x_{t-1}, y_{t-1})$ when a detected a kitchen knife at previous time $t-1$, the distance between two points is l_t. Here, if the distance l_t is less than the threshold th, the kitchen knife is not moving, and if the distance l_t is more than the threshold th, it is moving (we define that the user has the kitchen knife). The calculation method shows equation (1) and equation (2).

$$l_t = \|\boldsymbol{G_t} - \boldsymbol{G_{t-1}}\| = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2} \quad (1)$$

$$\begin{cases} \text{Holding a kitchen knife} & (lt \geq th) \\ \text{Not holding a kitchen knife} & (lt < th) \end{cases} \quad (2)$$

Fig. 4 shows the judgment of the kitchen knife holding on the hand.

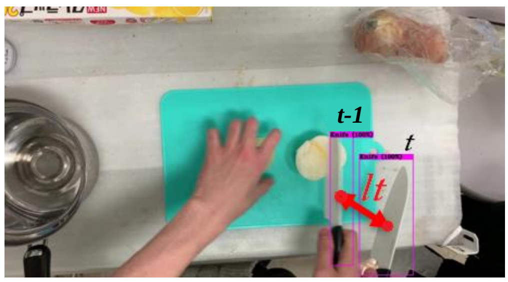


Fig .4. The judgment of the kitchen knife holding on the hand.

3) Is the distance between the kitchen knife and the hand without a kitchen knife?

The user holds the kitchen knife while cutting the food ingredients. Furthermore, the hand without a kitchen knife holds down the food ingredients. In such a situation, if the distance between the kitchen knife and the hand without the kitchen knife is large, the user is less likely to get injured; therefore, we decide the attention state from a distance of a perpendicular bisector that is calculated by the left line of the bounding box obtained by the YOLO algorithm and the fingertip coordinate of the hand without the kitchen knife. The coordinate of hand without the kitchen knife is obtained by the Open Pose algorithm [20]. Here, equation (3) shows that can calculate a straight line L passing between two points from the two coordinate points $P_1(x_1, y_1), P_2(x_2, y_2)$ of the bounding box.

$$L: (x_2 - x_1)(y - y_1) = (y_2 - y_1)(x - x_1) \qquad (3)$$

Next, we do deformation of equation (3), calculate a minimum distance between the coordinate points $\boldsymbol{H} = \{(Hx_1, Hy_1), (Hx_2, Hy_2), \cdots, (Hx_n, Hy_n)\}$ of hand without the kitchen knife and a straight line $L': ax + by + c = 0$ as shown in equation (4).

$$dist = \underset{i=1,\cdots,21}{\arg\min}\left(|aHx_i + bHy_i + c| / \sqrt{a^2 + b^2}\right) \qquad (4)$$

where, a, b and c are constant number.

Fig. 5 shows the calculation method of the distance between a kitchen knife and the fingertip of the hand without the kitchen knife.

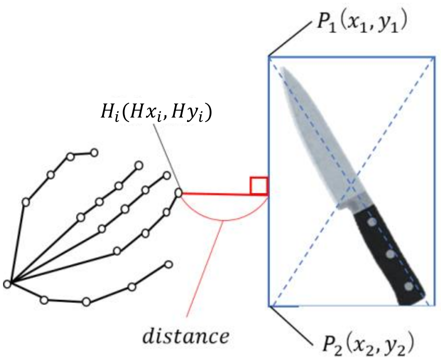


Fig. 5. The calculation method of the distance between a kitchen knife and the fingertip of the hand without the kitchen knife.

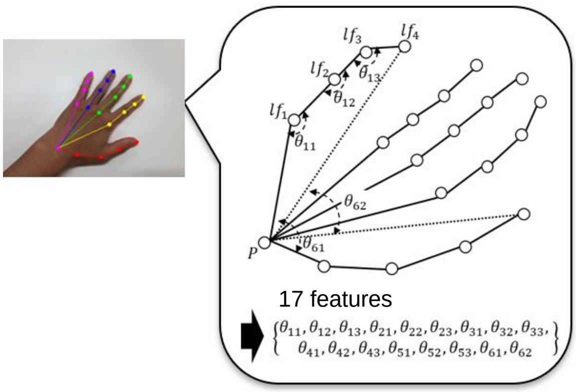


Fig. 6. The features of the hand skeleton information.

4) Are your hand without a kitchen knife open?

We get skeleton information of hand without the kitchen knife by the Open Pose algorithm to classify of opened state or closed state of hand without the kitchen knife. Then, this study defines 17 features that degrees between each finger joint ($\theta_{11} \sim \theta_{54}$), the degree between fingertips of thumb and little finger (θ_{61}), and the degree between fingertips of the base of the thumb and the base of the little finger (θ_{62}). Fig. 6 shows the detailed features of the hand skeleton information. We conducted pre experiments, could classify the opened state or the closed state hand that the accuracy was 82 percent [21].

IV. Experiments

As an experiment, we set a camera at a top in the lab's kitchen space, as shown in Fig. 7, to take a cooking video, extract only the cutting scene. Participants include a right-handed and a left-handed; therefore, we flipped a video of the left-handed participants, process as all participants right-hended. This study was conducted based on a review of the Research Ethics Review Committee at Ritsumeikan University (No. Kinugasa-Hito-2019-28).

Moreover, the dangerous situation is important to be estimated in advance rather than detected after they occur; therefore, we estimated attention state using long short term memory (LSTM). Specifically, the attention states are calculated using the results by experiment 1 and estimate the dynamic attention state, where, the learning data is the attention state of 0 to 10 seconds, the learning epoch was one thousand, the error was 0.083.

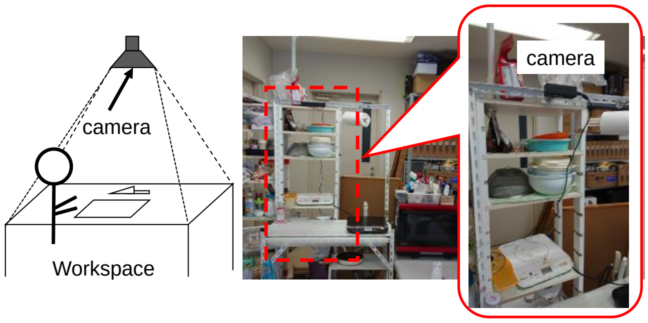


Fig. 7. An experiment's environment.

V. Results and Discussions

Fig. 8 shows the result of the dynamic attention state during cutting and Fig. 9 shows the confusion matrix of the result of the estimated attention state in all video. This time, we looked at the target video and annotate the video with the attention states.

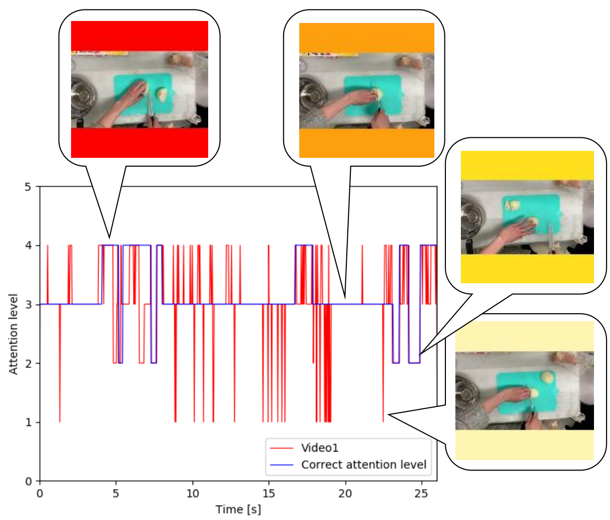


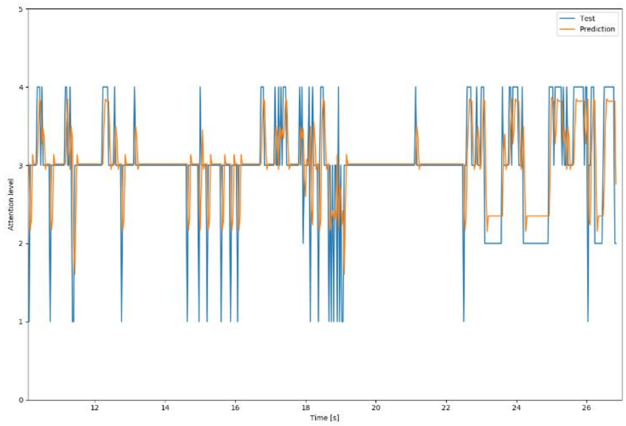
Fig. 8. An estimation result of the attention state.

As a result, the average accuracy of the estimated attention state was 74.3 percent. This cause is because attention state 3 and attention state 4 misrecognized each other. As a reason for misrecognized, the hand without the kitchen knife was sideways, so the hand skeleton information is not recognized by the Open Pose algorithm. Then, the recognition result of the attention state 3 and the attention state 4 is the attention state 1, which is the wrong result. We think our system misrecognized the attention state because the kitchen knife movement stopped when the ingredients were cut.

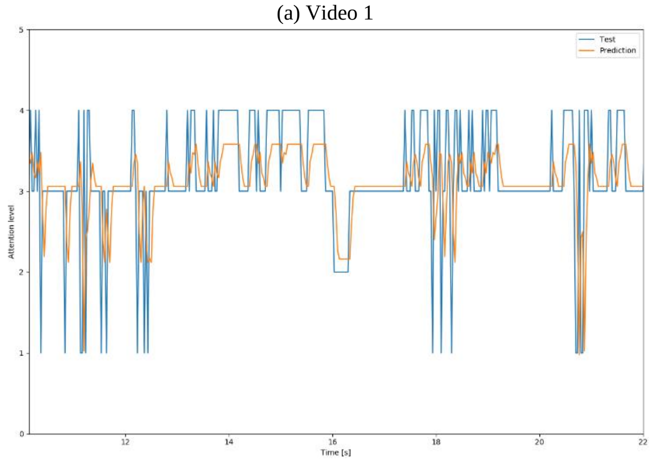
		Predcted label				
Attention states		0	1	2	3	4
True label	attention state 0	0.00	0.00	0.00	0.00	0.00
	attention state 1	0.00	0.00	0.00	0.00	0.00
	attention state 2	0.00	0.01	0.79	0.07	0.11
	attention state 3	0.00	0.04	0.02	0.72	0.22
	attention state 4	0.00	0.03	0.03	0.23	0.70
Average accuracy		***0.74***				

Fig. 9. Confusion matrix of estimation result of attention states.

Next, Fig. 10 shows the estimated results of the attention state by LSTM against two videos. As a result, we think that our system can estimate the attention state after 10 seconds.



(a) Video 1



(b) Video 2

Fig. 10. An estimation of attention state by LSTM.

VI. Conclusion

In this study, we focused a cognitive rehabilitation with cooking that was effective in improving cognitive function, patients and helper ware to decrease burdens, we aimed dynamic to grasp an attention state during cooking.

As the evaluation of attention states, we defined five stages from the attention states zero to four from four conditions that their conditions are "Is there a kitchen knife in the workspace?" "Do you have a kitchen knife?" "Is the distance between the kitchen knife and the hand without a kitchen knife?" and "Are your hand without a kitchen knife open?" respectively.

As a result, the average accuracy of the estimated attention state was 74.3 percent. Although the average accuracy of attention states was low, we think one of the causes was that when the hand's skeletal coordinates were not detected, the angle related to those coordinates was set to zero. Next, the classification of attention states was an average accuracy of 0.72. We believe this was of low accuracy because there were scenes where the Open Pose algorithm could not detect the hand angle and skeletal coordinates. We think the accuracy can improve by estimating the undetected skeletal information from past information to solve these problems.

Then, the experiments of this study focused on the scene during cooking with the cutting; however, there are scenes to be careful about in the cooking scene, such as using fire and placing cooking utensils. In future research, I would like to proceed with research targeting other scenes. Furthermore, we will develop a system that can quickly detect a user's abnormal state by accumulating daily data. In addition, I will add the features of head movement from using an egocentric vision, will detect whether the user is focused on the work. We will also study to add the environmental state of the workspace.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Sho Ooi and Kazuki Hashimoto conducted the research and developed the system, and wrote the paper; Sho Ooi, Kazuki Hashimoto, Haruo Noma and Mutsuo Sano discussed the system design and the evaluation method; all authors had approved the final version.

Acknowledgment

I am deeply grateful to Dr. Tabuchi, Dr. Saito, and Prof. Umeda of Keio University for the evaluation of attention states. This work was supported by JSPS KAKENHI Grant Numbers JP 19K20750.

References

[1] I. Itadume, "Disorder psychology of acurired brain injury," *Gakubunsha in Japan*, 2007.

[2] T. Honda, "The erhabilitation of the patient with acurired brain injury," *Igaku-Shoin in Japan*, 2006.

[3] Tokyo Metropolitan Government Brain Dysfunction Investigation Committee: Survey Report on the Actual Situation of Higher Brain Dysfunction in Tokyo, 2008.

[4] S. Ooi, M. Sano, S. Shibutani *et al.*, "Cognitive rehabilitation based on cooking-activities reflection support system for the independence of the patients with acquired brain injury," *Cognitive Rehabilitation*, vol. 20, no. 1, pp. 51-61, 2015.

[5] S. Tachigami, Rusk Institute of Rehabilitation Medicine, Brain Injury Day Treatment Program, Igaku-Syoin Ltd., Tokyo, 2010.

[6] M. Kato, "The development and standardization of Clinical Assessment for Attention (CAT) and Clinical Assessment for Spontaneity (CAS)," *Higher Brain Function Research*, vol. 26, no. 3, pp. 310-319, 2006.

[7] E. C. Brown, A. Casey, R. I. Fisch, and C. Neuringer, "Trail making test as a screening device for the detection of brain damage," *Journal of Consulting Psychology*, vol. 22, no. 6, p. 469, 1958.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 779-788.
[9] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “Realtime multi-person 2D pose estimation using part affinity fields,” in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
[10] S. Ooi, T. Ikegaya, M. Sano, H. Tabuchi, F. Saito, and S. Umeda, “Attention behavior evaluation during daily living based on egocentric vision,” *Journal of Advances in Information Technology*, vol. 8, no. 2, pp. 67-73, 2017.
[11] S. Ooi, M. Sano, H. Tabuchi, F. Saito, and S. Umeda, “Sustained Attention function evaluation during cooking based on egocentric vision,” in *Proc. the 19th IEEE International Symposium on Multimedia (ISM2017)*, 2017, pp. 107-113.
[12] B. Ji, Z. Wang, K. Xu and H. Liu, “The feasibility of a virtual reality system for attention analysis,” in *Proc. 2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI)*, 2019, pp. 207-211.
[13] O. Hikosaka, *Neural Mechanisms of Attention, Equiibrium Res*, vol. 57, pp. 543-549, 1998.
[14] M. Kato and H. Kashima, “Psychiatric clinical lumiere for specialist 10th attention disorders,” *Nakayama Shoten*, Tokyo, pp. 20-26, 2009.
[15] H. Kashima, T. Handa, M. Kato *et al*., “Attention disorder and frontal lobe injury,” *Neurological Progress*, vol. 30, pp. 847-858, 1986.
[16] M. M. Sohlberg and C. A. Mateer, *Attention Process Training. Association for Neuropsychological Research and Development*, Washington D. C., 1986.
[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788.
[18] G. Ross *et al.*, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
[19] S. Q. Ren *et al*., “Faster R-CNN: Towards real-time object detection with region proposal networks,” *Advances in Neural Information Processing Systems*, 2015.
[20] Z. Cao, T. Simon, S. E. Wei, and Y. Sheikh, “Realtime multi-person 2d pose estimation using part affinity fields,” in *Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
[21] K. Hashimoto, S. Ooi, H. Noma, M. Sano, H. Tabuchi, F. Saito, and S. Umeda, “A study on state detection baded on hand posuture for cognitive rehabilitation,” in *Proc. 6th Fall Meeting of Federation of Imageing Societies*, 2019.

Sho Ooi was born in Osaka, Japan. He received his Ph.D. degree in information science from Osaka Institute of Technology in 2018. Currently, he is an assistant professor with the Faculty of Information Science and Engineering in Ritsumeikan University, Japan. His research interests include computer vision, cognitive science, pattern recognition, and education technology. He is a member of the Information Processing Society of Japan (IPSJ), Institute of Electronics, Information, and Communication Engineers (IEICE), Institute of Image Electronics Engineers of Japan (IIEEJ), Robot Society of Japan (RSJ), and IEEE. He awarded the Best Student Paper Award in The 7th IIAE International Conference on Intelligent Systems and Image Processing, Partner Robot Challenge (Virtual) Service Robot Challenge (Virtual) WRS Special Award in World Robot Summit 2018, “Best in Navigation” in Robocup Committee.

Kazuki Hashimoto was born Aichi, Japan. He received his bachelor degree in engineering from Ritsumeikan University in 2019. His research interests include computer vision, deep learning, and cognitive science. He is a member of Institute of Image Electronics Engineers of Japan (IIEEJ).

Haruo Noma was born Ehime, Japan. He received his Ph.D. degree in engineering from Tsukuba University in 1994. Currently, he is a professor with the Faculty of Information Science and Engineering in Ritsumeikan University, Japan. His research interests include virtual reality, haptic interface, human interaction, cognitive science, pattern recognition, and education technology. He is a member of The Society of Instrument and Control Engineers (SICE), Human Interface Society (HIS), The Virtual Reality Society of Japan (VRSJ), Robot Society of Japan (RSJ), Institute of Electronics, Information, and Communication Engineers (IEICE), ACM, and IEEE.

Mutsuo Sano was born Ehime, Japan. He received his Ph.D. degree in engineering from Kyoto Institute of Technology in 1981. Currently, he is a professor with the Faculty of Information Science and Technology in Osaka Institute of Technology, Japan. His research interests include computer vision, cognitive science, cooking media, cooking support, and human robot communication. He is a member of Information Processing Society of Japan (IPSJ), the Institute of Electronics, Information and Communication Engineers (IEICE), The Institute of Image Electronics Engineers of Japan (IIEEJ), Robot Society of Japan (RSJ), The Japanese Society for Artificial Intelligence (JSAI), and IEEE.