DEIM Forum 2021 J24-3

Huu-Long PHAM†, Yusuke YAMAMOTO††, Takehiro YAMAMOTO†††, and Hiroaki OHSHIMA†,†††

† Graduate School of Applied Informatics, University of Hyogo
7-1-28 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650–0047
††† School of Social Information Science, University of Hyogo
8-2-1 Gakuennishi-machi, Nishi-ku, Kobe, Hyogo 651–2197
†† Faculty of Informatics, Shizuoka University
3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka 432–8011
E-mail: †{aa19e510,ohshima}@ai.u-hyogo.ac.jp, ††yamamoto@inf.shizuoka.ac.jp,
†††t.yamamoto@sis.u-hyogo.ac.jp

Abstract In this research, we developed a system that helps information workers improving concentration level and productivity at work using distractions detection with wearable sensor device. Nowadays, information workers are facing a lot of distractions at work, including both internal distractions and external distractions. Those distractions may lead to the reduction of concentration, which is one main reason of the decrease in term of productivity. Our system uses a wearable device and a desktop application for distraction detection. Distractions such as accesses to work-unrelated websites are detected by the desktop application, while physical distractions from user like sleeping or picking up and using smartphones are detected using sensor data from the wearable device. We constructed a machine learning model to recognize user's activities and use recognized activity information to detect distractions occurred. In this research, we introduce three methods for concentration level improvement using detected distractions, which are: distraction notification, concentration bar and concentration competition.

Key words distractions; concentration; interruptions; multitasking; productivity; human activity recognition

1 Introduction

Information workers experience many kinds of distractions while working. Those distractions can be either external distractions, like incoming calls, or internal distractions, like picking up and using smartphones unconsciously. Mark et al. [13], through a workplace experiment, showed that focus time duration of information work is fairly short, with only a median span of 40 seconds. In a field study of the multitasking behavior of computer users [9], Isqbal et. al found that every hour, an information worker's primary tasks are interrupted by an average of 4.28 email alerts and 3.21 other messages alerts, which averages 3.74/hour overall. They also showed that the average time to return to the suspended task was 9 minutes and 33 seconds for emails, and 8 minutes for other messages. In addition, Gonzalez et. al [7] found that activities in the workplace including online work and interactions with other people shift every three minutes on average.

In some cases, information workers can benefit from distractions and task switching in term of work load reduction. However, since human's intentional resources is finite [21], those distractions in most cases can cause concentration level reduction and distract people from the current primary task. Researches showed that switching different tasks results in a 50% longer time required to finish those tasks in total, compared to focus on each task through completion [5]. A study at Salary.com in 2004 showed that over 750 participants, 89% reported that they are wasting at least 30 minutes at work daily. Since in many cases information workers usually fall into distractions undeliberately, it is necessary to notice people when those distractions happened, or blocking people from distractions completely. While there are many commercial tools that can block distractions such as social websites completely [12], prevention from distractions that performed internally by human like picking up smartphones without realizing, standing up going somewhere, is not yet developed.

In this research, we developed a concentration level and productivity improvement system. Our system includes a wearable devics (Figure 1) and a desktop application. Desktop application records user's websites browsing data, and sends a notification whenever accesses to work-unrelated websites like Youtube.com, Facebook.com, etc are detected. Alerts are sent to user through the wearable device's speakers

Figure 1 Self-monitoring device

to draw user's attention back from distractions. Our system can not only detect distractions that occurred on user's computer, but also can detect physical distractions which were performed by user, like picking up phone by a behavior without realizing, or standing up going to grab some drinks, or even sleeping. In order to detect those distractions, we developed a machine learning model, which uses sensor data from the wearable device to recognize user's current activities. The main contributions of our work is summarized in the following:

- We developed a distractions detection system using human activity recognition from wearable device and user's web browsing data. User's physical activity information is used in order to detect distractions such as using smartphone unconsciously or sleeping.
- We proposed interactive methods using a wearable device that helps user improving their concentration level and productivity using information from detected distractions. Details of these methods are explained in section 5.

2 Related work

2.1 Concentration and Distractions

Basically, concentration level of a information worker during cognitive tasks is measured by activities in the prefrontal cortex (PFC), such as blood flow or change in oxygenated hemoglobin (oxy-Hb). Measurements of those activities required devices like functional near-inferred spectroscopy (fNIRS) or functional magnetic resonance imaging (fMRI) [3], which are very costly. There are also researches that proposed concentration level measurements methods in less expensive ways, like by using Smart Eyewear [10], or by using an intellectual concentration index [16].

Although high concentration level is ideal, it is difficult for information users to keep focusing on their ongoing tasks at work. Mark et. al [13] showed that focus time duration of information work last only 40 seconds on average. While information workers often distract themselves unconsciously, there are several causes of distractions, such as organizational environment, individual differences or usage of mobile devices [4] [12].

2.2 Human activity recognition using wearable sensor data

Human Activity Recognition using wearable sensor data is a popular task nowadays. Data from sensors like accelerometer and gyroscope are often used in these tasks [11] [20] [2] [8] [17], since such sensors can cover information about user's posture, which is useful to determine his/her current activities. Features selection and Features generation is one of the most difficult challenges in human activity recognition tasks. Usually, input features are extracted manually from raw sensor data [11] [20]. Although this approach has proven to be effective, it faces to big challenges:

(1) First, sensor data measurements are noisy. Measured sensor data are not unified even same type of devices were used.

(2) Second, it is difficult to find the most useful features to accommodate noise patterns in sensor data and user behaviors.

Because of the heterogeneity in sensor data measurement and user behaviors, performances of activity recognition models in real world are often less accurate than which they were reported in papers [18].

Later researches proposed the usage of Convolution Neural Network and Recurrent Neural Network such as LSTM in order to either extract useful features automatically or exploit the temporal relations between sensors data at different time step [8] [17] [15]. Especially, there are several researches proposed the combination of both CNN and RNN [23] [22] [1].

3 Concentration level improvement system

In this research, we developed a system that improve concentration and productivity level by detecting distractions. Target users are people who work with computers.

3.1 System overview

Our system consists of the following elements:

(1) **Self-monitoring device:** Watch-type wearable device made with M5 Stack Fire(Note1). We use six types of sensor data from this device for user activity recognition. The device also has a LCD display and a speaker, which are used to interact with user.

(2) **Desktop application:** We developed a multiplatform desktop application that can run on several OS including Windows, MacOS and Linux. Desktop application logs user's websites browsing data and alert him/her when ac-

(Note1): https://m5stack.com/products/fire-iot-development-kit

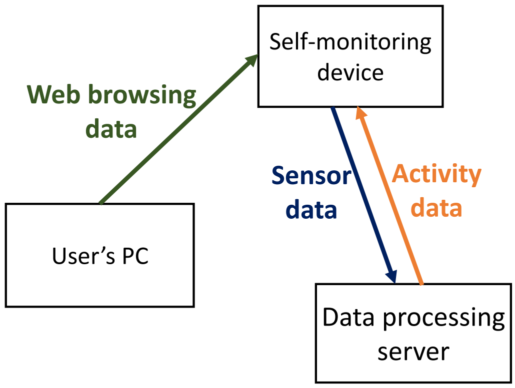

Figure 2　System's data processing

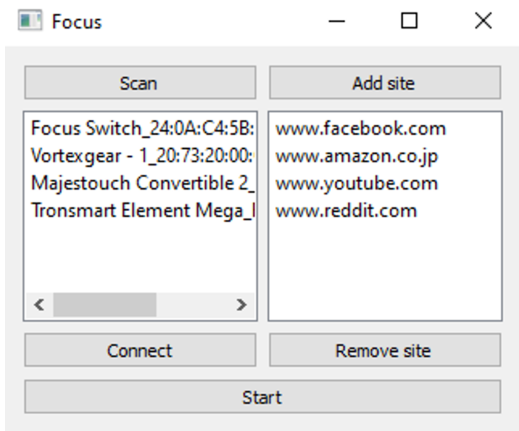

Figure 3　Desktop Application. From right side of the application, user can choose which site to be alerted when access to that site was detected. Nearby BLE devices are listed in the left side after clicking Scan button. A connection between Self-monitoring device and Desktop Application is established by clicking Connect button. The application starts logging user's web access data after Start button is clicked.

cesses to distractive work-unrelated websites were detected.

(3) **Data processing server:** We constructed a server that uses encrypted MQTT protocol to receive sensor data from Self-monitoring device and recognizes activities performed by user using received sensor data.

Figure 2 illustrates how data is processed in our system. User's websites browsing data recorded by desktop application is sent from user's PC to the Self-monitoring device through BLE(Bluetooth Low Energy) [6]. Sensor data from Self-monitoring device is sent to data processing server using encrypted MQTT protocol. Server uses sensor data to recognize user's activity, and sends activity information back to Self-monitoring device

3.2 Self-monitoring device

In this study, we used M5 Stack Fire module to develop a wearable device called Self-monitoring device. Figure 1 shows picture of the device. The device have following main features:

(1) **Timer:** Researches showed that using a timer can result in improvement of concentration level. Our device can be used a timer to determine how long next task is going to last. User can check remaining time of the current task is then alerted when the timer goes to zero.

(2) **Sensor data recorder:** There are 6 type of sensors inside the device: 3-axis Accelerometer sensor and 3-axis Gyroscope sensor. Sensor data is collected at a frequency of 50Hz, which means 50 times per second. For every one second, sensor data is sent to server and used to recognize user's physical activities.

(3) **Notification and alert:** Whenever distractions are detected, users are then alerted to go back to their ongoing task.

(4) **Concentration visualization:** Concentration level is visualised and shown to user through the device's display.

3.3 Website browsing data tracking desktop application

We used Python to develop a multiplatform desktop application that runs on different OS including Windows, MacOS and Linux. Main feature of this application is: **Logging user's web browsing data**. Whenever accesses to distractive work-unrelated websites like Facebook, YouTube, Reddit, etc are detected, user is then alerted by Self-monitoring device to close those sites and go back to the current task. Data communication between Self-monitoring device and Desktop application is implemented through BLE (Bluetooth Low Energy) protocol. User can choose which sites to be alerted from the application. Desktop application's interface is illustrated in figure 3. Usages of the desktop application are listed step by step as below:

Step 1: Open the application.

Step 2: Use "Add site" and "Remove site" buttons to choose which sites whose accesses should be observed.

Step 3: Click "Scan" button to get nearby BLE devices.

Step 4: Choose the appropriate Self-monitoring device and click on "Connect" button to establish a connection between Desktop Application and the chosen Self-monitoring device.

Step 5: Click on "Start" button to start the application.

4 Human Activity Recognition

In this research, we use sensor data from the Self-monitoring device to recognize user's activity. This section describes methods used for human activity recognition.

4.1 Problem Definition

We use **one second** of sensor data to recognize user's physical activities occurred in that one second. Because it is possible that multiple activities were performed simultaneously or sequentially, human activity recognition in this research is solved as a **multi labels** problem. For example, in one second, user could be walking and using smartphone at the same time. Nevertheless, because one second is considerably short, we expect that only one or two activities are performed in one time span.

4. 2 Recognized Activities

Human activities that are recognized in this research are listed as below:

(1) Working
(2) Walking
(3) Standing
(4) Using smartphone
(5) Sleeping
(6) Reading document
(7) Reading book
(8) Other

"Using smartphone", "Walking", "Sleeping" and "Other" are listed as **distractions**. Since breaks are necessary during while working, user is not immediately alerted when distraction was recognized. When one ore more distractions are recognized to be continuously proceeded for over five minutes, users are then alerted to go back to their current tasks.

4. 3 Human activity recognition using Bidirectional LSTM and Transformer's Encoder

In this research, we recognize human activity from sensor data using two feature extraction methods.

(1) **Bidirectional LSTM:** With time series data, temporal relation between each time-step is a important information, and often extracted manually in previous researches, like calculating means and variances. By using recurrent neural network like LSTM, we expect that useful temporal relation between timesteps can be automatically exploited more correctly. We used Bidirectional LSTM since human activity recognition problem in this research does not require real time recognition.

(2) **Transformer's Encoder:** In practice, with long sequence input, LSTM models can suffer from vanishing gradient problem. Therefore, in some cases, output of the later timesteps in the sequence may not have enough information from earlier the timesteps. Transformer structure, which was developed by Vaswani et al. [19], use self-attention technique to solve this problem. In this research, beside the Bidirectional LSTM method, we also use a stack of 6 Transformer's Encoders for feature extraction.

Data from Accelerometer and Gyroscope from wearable device are used for human activity recognition. Because each sensor has three axes x, y and z, there are a total of six types of sensor data. Sensor data is captured with a frequency of 50Hz, which means 50 times per second. Since there six types of sensor data, both Bidirectional LSTM model and Transformer's Encoder model take a a 6 X 50 matrix as input, where each column of the matrix is the sensor data recorded

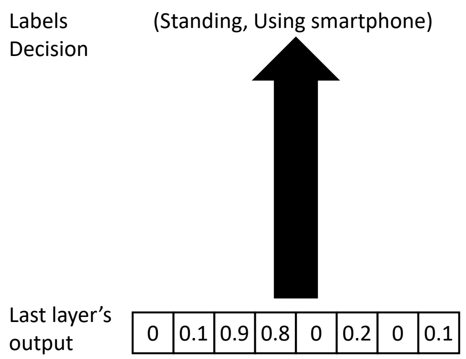

Figure 4 Activity Labels Decision Example. Since value of 3^{rd} and 4^{th} dimensions are larger than 0.5 (0.9 and 0.8), the two corresponding labels "Standing" and "Using smartphone" are chosen.

in one time step. Raw sensor data is normalized using the following equation:

$$z_k^t = \frac{(x_k^t - u_k)}{s_k}$$

Raw sensor data x_k^t at k channel and time-step t is subtract by mean u_k, and then divided by variance s_k. u_k and s_k are calculated using training data.

Both Bidirectional LSTM model and Transformer's Encoder model outputs a vector with eight dimensions, which is the same number of recognized activities as described in section 4. 2. Sigmoid function is applied on each dimension of the output vector. Therefore, each dimension's value ranges from 0 to 1, and is used as probability for a particular activity. Output result decision based on output vector is illustrated in figure 4. For training, we used Binary Cross-Entropy Loss function to to calculate loss between model's output and true activity labels. Adam Optimization was used for training with an initial learning rate of 10^{-3}, decreased by two times every 10 epochs. We trained both model for 100 epochs and chose the epoch with lowest validation loss.

4. 3. 1 Bidirectional LSTM model

Bidirectional LSTM model's structure is illustrated in Figure 5. We use one forward and one backward LSTM layer. Both layer outputs a 64 dimensions vector for each timestep. Output of forward layer and backward layer for each timestep is concatenated into a vector with 128 dimensions. After that, timestep-wise concatenation is also performed, which results a $50 * 128 = 6400$ dimensions vector. This vector is fed into a fully connected layer to output a vector with 8 dimensions. Batch normalization and sigmoid are then applied to generate final output.

4. 3. 2 Transformer's Encoder model

Transformer's Encoder model's structure is illustrated in

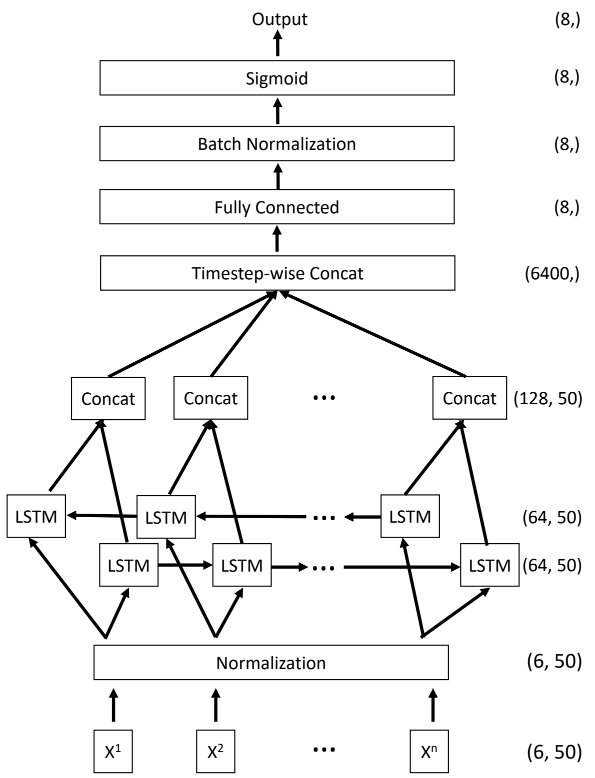

Figure 5 Structure of Bidirectional LSTM model

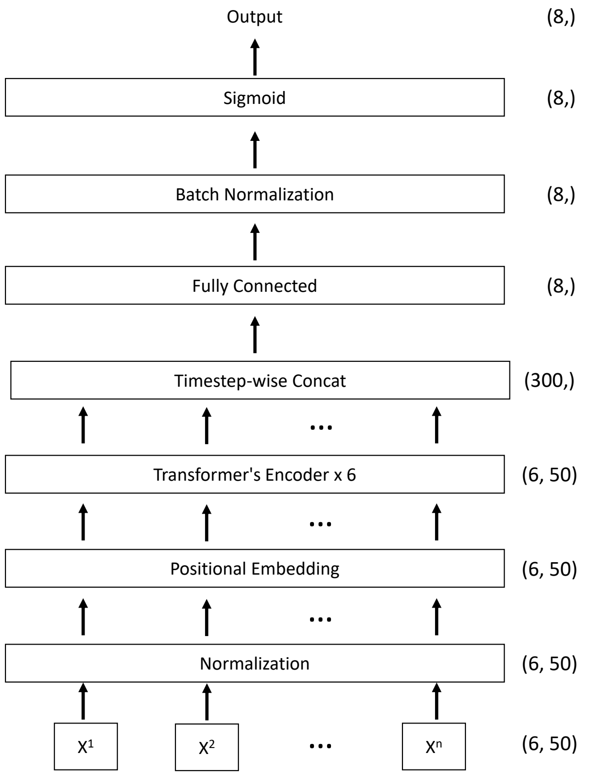

Figure 6 Structure of Transformer's Encoder model

figure 6. Raw input data after normalized is feed into a Positional Embedding layer. We used the sine function from the original transformer paper [19] in order to embed normalized sensor data at every timestep and dimension:

Table 1 Accuracy of Bidirectional LSTM model and Transformer's Encoder model

Activity	Bi-LSTM	Transformer's Encoder
Working	82.82%	**87.97%**
Walking	96.46%	**96.58%**
Standing	**97.00%**	96.53%
Using smartphone	**84.74%**	83.87%
Sleeping	**86.39%**	84.53%
Reading document	**87.07%**	84.41%
Reading book	**88.84%**	87.27%
Other	**88.81%**	85.96%
Average	**89.02%**	88.39%

$$PE_{(pos,2i)} = sin(pos/10000^{2i/6})$$

where pos is the timestep position and i is the sensor dimension. Since the window size is 50 and there are 6 types of sensor data, pos ranges from 0 to 49 and i ranges from 0 to 5. In this research, we use a stack of 6 Transformer's Encoder layers with 8 multi attention heads. The dimension of feed-forward layer for each Encoder layer is 512. Dropout with a rate of 0.3 is also used for each Encoder layer. Output of the Transformer's Encoders stack remains the same dimension with input data, and is concatenated timestep wisely into a $50 * 6 = 300$ dimensions vector. This vector is fed into a fully connected layer to output a 8 dimensions vector. Batch normalization and sigmoid are applied as in the Bidirectional LSTM model to generate final output.

4.4 Human activity recognition evaluation

We constructed a dataset which has data of ten people. Sensor data of each activity was recorded in 10 minutes for every individual. Since there are 8 activities, the total duration of the dataset is: $10 * 8 * 10 = 800$ minutes. Evaluation was implemented using Leave One Group Out cross validation. In every round, data of one person is used for validation and the remaining data is used for training.

Table 1 shows accuracy of each model. Bidirectional LSTM model performs slightly better.

5 Concentration level improvement using distraction detection

In this research, we detect distractions occurred while user is working and use that information to improve concentration level. This section describes the methods were used for concentration level improvement.

5.1 Distraction detection

Our system is able to detect two kinds of distractions: **accesses to work-unrelated websites** and **physical activities that are considered as distractions**. These two kinds of distractions are described as follow:

(1) **Accesses to work-unrelated websites:** While working with computers, information worker sometimes checks SNS or reads news unconsciously. Although watching those sites can benefit in term of workload reduction, doing so frequently can reduces focus level and productivity. Our system uses a desktop application to capture user's websites browsing activities, and alert him/her when accesses to distractive websites were detected.

(2) **User's physical activities:** Activities like sleeping or using smartphone unconsciously are considered as distractions. User often falls in these distractions by a behavior, without even realizing. Researches show that recovering and reconstructing from interruptions can require a lot of time, depends on the duration of interrupt event. Our system uses sensor data from Self-monitoring device to recognize user's activities and extract distractions from recognized activity information.

5.2 Concentration level improvement methods

We propose three methods to improve concentration level based on detected user's distractions:

(1) **Distraction notification:** User is alerted by Self-monitoring device whenever distractions were detected. However, considering that short breaks are important, notifications are sent to user unless distractions were lasting for 5 minutes continuously.

(2) **Concentration bar:** Used to visualize user's concentration. Concentration is calculated using the following equation:

$$\text{concentration} = \frac{T_{\text{focused}}}{T_{\text{total}}}$$

where T_{focused} is the seconds that user was focusing on the current task, and T_{total} is the total seconds have passed. In this research, distraction detection is performed for every one second. If no distraction was detected, our system judges that user is focusing on the current task. Two types of concentration are visualized: Long term concentration and Short term concentration. Long term concentration is calculated using data from start of the task, while Short term concentration is calculated using only data from nearest 30 seconds. Two type of concentration bars are illustrated in Figure 7.

(3) **Concentration competition:** Comparison and competition can help to improve motivation. Total focused time of every user is saved in a database, and is used to find out top users, daily and weekly. Focused times of top users are displayed on the Self-monitoring device's screen, along with data of the current user as illustrated in Figure 7.

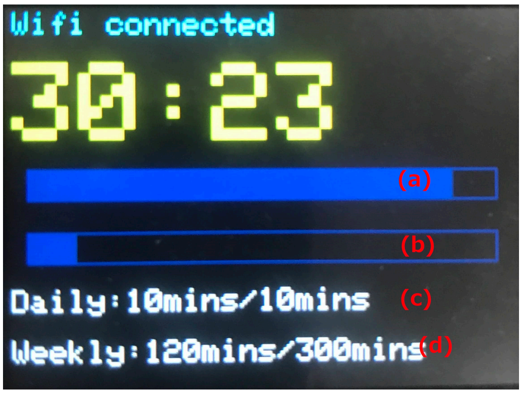

Figure 7 Illustration of concentration bar and concentration comparison. (a) and (b) are concentration bars for Long Term Concentration and Short Term Concentration, respectively. (c) and (d) shows daily and weekly total focused minutes of current user and top user. Total focused minutes of current user is placed before the slash while top user is placed after.

6 Experiment

We conducted an experiment with the cooperation of eight participants, who are all undergraduate students in the Informatic field. The goal of this experiment is to indicate the effectiveness of our system in concentration level and productivity improvement. This section describes details and results of the experiment.

6.1 Experiment details

Participants were divided equally into two groups, called Group A and Group B. Each group has two males and two females. The experiment's task is to read the second chapter of a book called "Introduction To Machine Learning With Python" [14], and summarize contents into a slide using Microsoft PowerPoint. Participants were asked to read and summarize contents simultaneously. Every participant executes the task in sixty-minutes, for two turns. Participants from Group A use our system in the first turn only, while participants from Group B use our system in the second turn only. The experiment takes a total of two hours and fifty minutes, including warming up, breaks, and task executions. Experiment's overall process for each group is illustrated in Figure 8.

Participants were asked several questions after each turn. The questions are:

(1) Task question 1 (TQ-1): Could you focus during the task?

(2) Task question 2 (TQ-2): Do you think your slide is good in term of "quantity"?

(3) Task question 3 (TQ-3): Do you think your slide is good in term of "quality"?

(4) Task question 4 (TQ-4): How many pages of the book you was able to read?

(5) Task question 5 (TQ-5): Do you satisfy with the task you have just executed?

Answers for each question range from 1 to 4, where

Table 2　Participant's responses for all task questions, overall question 1 and 2

id	sex	group	with system					without system					OQ-1	OQ-2
			TQ-1	TQ-2	TQ-3	TQ-4	TQ-5	TQ-1	TQ-2	TQ-3	TQ-4	TQ-5		
1	F	A	3.00	3.00	2.00	12.10	4.00	3.00	4.00	2.00	18.20	3.00	2.00	3.00
2	M	A	2.00	2.00	2.00	18.00	2.00	2.00	2.00	3.00	26.00	3.00	4.00	3.00
3	F	A	3.00	3.00	2.00	18.00	3.00	2.00	1.00	2.00	14.00	2.00	2.00	2.00
4	M	A	3.00	3.00	2.00	89.00	3.00	2.00	2.00	3.00	89.00	2.00	3.00	3.00
5	M	B	4.00	4.00	4.00	53.00	4.00	4.00	1.00	3.00	44.00	3.00	3.00	4.00
6	F	B	3.00	4.00	2.00	38.00	3.00	3.00	3.00	2.00	22.00	3.00	3.00	3.00
7	M	B	3.00	3.00	3.00	18.40	4.00	3.00	2.00	2.00	4.70	3.00	3.00	3.00
8	F	B	2.00	2.00	4.00	16.00	3.00	4.00	2.00	4.00	3.50	3.00	1.00	2.00
Average			2.88	3.00	2.63	32.81	3.25	2.88	2.13	2.63	16.55	2.75	2.63	2.88

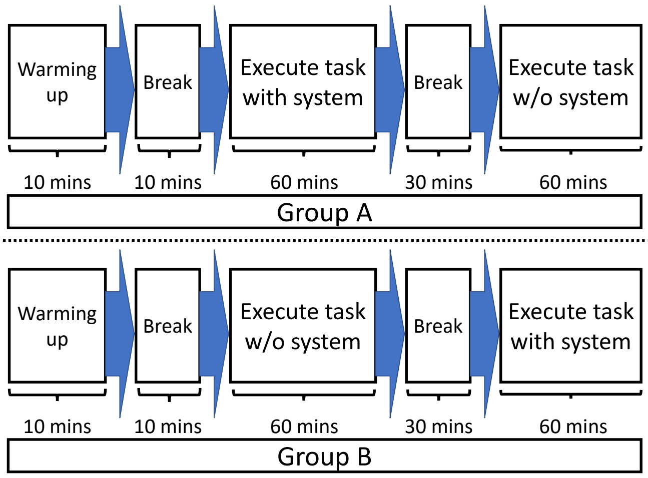

Figure 8　Experiment Process

1=disagree, 2=somewhat disagree, 3=somewhat agree and 4=agree. After finishing the experiment, participants were asked to answer more 5 questions, which are:

(1) Overall question 1 (OQ-1): Did you feel the effectiveness of the system?

(2) Overall question 2 (OQ-2): Do you want to use the system in the future?

(3) Overall question 3 (OQ-3): What is the pros of the system?

(4) Overall question 4 (OQ-4): What is the cons of the system?

(5) Overall question 5 (OQ-5): What features do you want the system to have?

Overall question 1 and 2 has four answers which are same with answers of task questions, while overall question 3, 4, 5 have text answers.

6.2 Experiment results

Table 2 shows responses of participants for all task questions with and without system, overall question 1 and overall question 2. "TQ" stands for task question and "OQ" stands for overall question. According to responses of overall question 1 and 2, 5/8 participants could feel the effectiveness of the system, and 6/8 would want to use the system in the future.

In order to evaluate the system's effectiveness, we performed a paired samples t-test on responses of task questions. The hypothesis is: there are differences between executing task with system and executing task without system. However, there was no significant difference and null hypothesis was not rejected. Overall, participants feel that their concentration levels were there same whether using system or not. Self-evaluations of the slides in term of quality are roughly equal too. However, we notice that using the system can slightly improve quantity of the task. With our system, participants were able to read 32.81 pages on an average, compared to 16.55 pages without the appearance of the system. Self-evaluation of slide's quantity is also slightly higher with the present of the system. Distraction notification feature can be the reason of this slight difference. Since users take alerts from Self-monitoring device whenever distractions were detected, they do not fall deeply in distractions and are able to go back to the current task quickly.

System's pros which were reported by participants are as follows:

- Visualization helps knowing current concentration level.
- Competition with top user motivated.
- Timer helps managing time efficiently.
- Alerts from device while using smartphone made going back to current task.

Beside that, system's cons were also reported, such as:

- Alerts are noisy.
- Device's size is too big.
- Alerted by device while focusing.

We can see that while some users benefit from distraction notification, others think that alert are annoying and noisy. In addition, mistakes from distraction detection sometimes can even cause reduction of concentration level.

7 Conclusion

In this research, we proposed a method to improve Concentration level using distraction detection. Distractions like accesses to work-unrelated websites such as YouTube, Facebook are captured by a desktop application. Physical distractions such as using smartphone unconsciously, sleeping are detected using sensor data from a wearable sensor device. We constructed a machine learning model to recognize user's activities, and use activity information to detect distractions occurred. The human activity recognition model takes one second of sensor data from the wearable device as input and outputs user's activities with 91% accuracy. Our system attempts to improve user's concentration level by three methods: distraction notification, concentration bar and concentration competition. We conducted an experiment with the cooperation of eight participants. Although experiment's results showed that there was no significant improvement using our system in term of concentration level or work's quality, slight improvement in term of work's quantity was reported.

In the future, we plan to evaluate the system furthermore by evaluating slides created by participants. We also plan to try different window size for human activity recognition, like two seconds, three seconds, etc. Finally, we also plan to exploit other effective methods for concentration level and productivity improvement.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Numbers JP18H03243, JP17H00762, JP18H03244, and by ROIS NII Open Collaborative Research 2020-20S1001.

References

[1] Waqar Ahmad, Bibi Misbah Kazmi, and Hazrat Ali. Human activity recognition using multi-head cnn followed by LSTM. In *Proceedings of ICET'19*, pp. 1–6, 2019.

[2] Akram Bayat, Marc Pomplun, and Duc Tran. A study on human activity recognition using accelerometer data from smartphones. *Procedia Computer Science*, Vol. 34, pp. 450–457, 2014.

[3] Mickael Causse, Zarrin Chua, Vsevolod Peysakhovich, Natalia Del Campo, and Nadine Matton. Mental workload and neural efficiency quantified in the prefrontal cortex using fNIRS. *Scientific Reports*, Vol. 7, pp. 1–15, 2017.

[4] Laura Dabbish and Robert E. Kraut. Controlling interruptions: Awareness displays and social motivation for coordination. In *Proceedings of CSCW'04*, pp. 182–191, 2004.

[5] Richard Gendreau. The new techno culture in the workplace and at home. *Journal of American Academy of Business*, Vol. 11, pp. 191–196, 2007.

[6] Carles Gomez, Joaquim Oller Bosch, and Josep Paradells. Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology. *Sensors*, Vol. 12, pp. 11734–53, 2012.

[7] Victor M. González and Gloria Mark. "Constant, constant, multi-tasking craziness": Managing multiple working spheres. In *Proceedings of CHI'04*, pp. 113–120, 2004.

[8] Sojeong Ha and Seungjin Choi. Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors. In *Proceedings of IJCNN'16*, pp. 381–388, 2016.

[9] Shamsi T. Iqbal and Eric Horvitz. Disruption and recovery of computing tasks: Field study, analysis, and directions. In *Proceedings of CHI'07*, pp. 677–686, 2007.

[10] Shoya Ishimaru, Kai Kunze, Katsuma Tanaka, Yuji Uema, Koichi Kise, and Masahiko Inami. Smart eyewear for interaction and activity recognition. In *Proceedings of CHI EA'15*, pp. 307–310, 2015.

[11] Adil Khan, Soo-Young Lee, and Tae-Seong Kim. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IIEEE Transactions on Information Technology in Biomedicine*, Vol. 14, pp. 1166–1172, 2010.

[12] Gloria Mark, Mary Czerwinski, and Shamsi T. Iqbal. Effects of individual differences in blocking workplace distractions. In *Proceedings of CHI'18*, pp. 1–12, 2018.

[13] Gloria Mark, Shamsi Iqbal, Mary Czerwinski, Paul Johns, and Akane Sano. Neurotics can't focus: An in situ study of online multitasking in the workplace. In *Proceedings of CHI'16*, pp. 1739–1744, 2016.

[14] Andreas C. Müller and Sarah Guido. *Introduction to Machine Learning with Python*. O'Reilly Media, Inc., 2016.

[15] Abdulmajid Murad and Jae-Young Pyun. Deep recurrent neural networks for human activity recognition. *Sensors*, Vol. 17, pp. 2556–2573, 2017.

[16] Fumiaki Obayashi, Hirotake Ishii, and Hiroshi PShimoda. Intellectual concentration index and concentration-improving lighting. *Panasonic Technical Journal*, Vol. 62, pp. 50–55, 2016.

[17] Madhuri Panwar, S Ram Dyuthi, K Chandra Prakash, Dwaipayan Biswas, Amit Acharyya, Koushik Maharatna, Arvind Gautam, and Ganesh R Naik. CNN based approach for activity recognition using a wrist-worn accelerometer. In *Proceedings of EMBC'17*, pp. 2438–2441, 2017.

[18] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne, and Mads Møller Jensen. Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition. In *Proceedings of SenSys'15*.

[19] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, undefinedukasz Kaiser, and Illia Polosukhin. Attention is All You Need. In *Proceedings of NIPS'17*, pp. 6000–6010, 2017.

[20] George Vavoulas., Charikleia Chatzaki., Thodoris Malliotakis., Matthew Pediaditis., and Manolis Tsiknakis. The mobiact dataset: Recognition of activities of daily living using smartphones. In *Proceedings of ICT4AWE'16*, pp. 143–151, 2016.

[21] Christopher Wickens. Multiple resources and mental workload. *Human factors*, Vol. 50, pp. 449–55, 2008.

[22] Kun Xia, Jianguang Huang, and Wang Hanyu. LSTM-CNN architecture for human activity recognition. *IEEE Access*, Vol. 8, pp. 56855–56866, 2020.

[23] Shuochao Yao, Shaohan Hu, Yiran Zhao, Aston Zhang, and Tarek Abdelzaher. Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In *Proceedings of WWW'17 Companion*, pp. 351–360, 2017.