Study of Measuring Motor Function in Dementia by Tapping Task Using IoT and Image Recognition

1st Sinan Chen
Kobe University,
1-1 Rokkodai-cho, Nada,
Kobe, 657-8501, Japan
chensinan@gold.kobe-u.ac.jp

2nd Atsuko Hayashi
Kobe University,
7-10-2 Tomogaoka, Suma,
Kobe, 654-0142, Japan
a-hayashi@pearl.kobe-u.ac.jp

3rd Masahide Nakamura [1,2]
[1]*Kobe University,*
[2]*RIKEN Center for Advanced Intelligence Project,*
Tokyo, 103-0027, Japan
masa-n@cs.kobe-u.ac.jp

2023 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD) | 979-8-3503-1335-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICTMOD59086.2023.10438119

***Abstract*—Dementia has emerged as a significant issue in contemporary Japanese society. Assessing finger motor function is pivotal for diagnosing and rehabilitating individuals with dementia. In this study, we design a web application to gauge manual dexterity through tapping exercises. Our fundamental concept combines finger identification using a web camera and tactile panel input to quantify finger actions. In our proposed methodology, we amalgamate real-time visuals captured by a web camera into the hand.js module within the MediaPipe framework, autonomously generating 21 critical finger data points. Concurrently, we establish touch panel control buttons on the web interface. In the context of the operating interface, participants manipulate their fingers in sync with numbers or features displayed every two seconds and utilize touch panel input to log button presses. This approach facilitates an all-encompassing assessment of finger motor aptitude.**

***Index Terms*—motor function, tapping task, finger dexterity, IoT, image recognition, Web service**

I. Introduction

In recent years, Japan has been experiencing a rapid aging population, leading to concerns about the increasing number of dementia patients. Dementia is a condition characterized by a decline in memory and cognitive functions, which hinders everyday life and social interactions. The number of dementia patients in Japan is projected to rise in the future. Against this societal backdrop, there has been a growing interest in healthcare for dementia. Dementia not only leads to a decline in cognitive functions but also affects motor functions. In particular, Alzheimer's disease is known to be associated with a noticeable decline in motor function [1]. Research has been conducted to diagnose dementia in its early stages by measuring changes in motor functions such as fine motor skills and walking abilities in dementia patients [2].

A method commonly employed to detect a decline in fine motor skills is the tapping task. The tapping task involves the experimenter providing instructions, often using numbers or other cues, and assessing whether the subject can move their fingers appropriately in response to these instructions. Traditional methods relied on providing instructions on paper or other physical mediums, with the experimenter visually assessing the results. However, this approach posed significant burdens on both the experimenter and the subjects. Efforts have been made to automate the evaluation of the tapping task using methods such as magnetic sensors or touch panels [3]. However, these methods have mainly focused on assessing only the first and second fingers, with limited attention given to evaluating all five fingers. Additionally, a challenge with these methods is that the instruction-giving aspect and the input aspect are independent, making it difficult to conduct evaluations and analyses within an application.

The goal of this study is to address these challenges by developing a web service for measuring fine motor skills using the tapping task. The proposed service automates the execution of the tapping task through simple interactions and records and evaluates the results automatically. Specifically, concerning hardware, we employ a USB fixed camera and a general-purpose computer. On the software front, we designed a web interface for conducting tapping tasks. The system corresponds to numerical and Japanese kana characters presented every two seconds, allowing it to record data of buttons pressed on the touch panel. Concurrently, utilizing a pre-trained model based on images from the USB camera, the system can also capture changes in finger recognition features.

Based on our proposed methodology, we proceed with system implementation. As a key technological aspect, we coordinate a web server built with *node.js* and a *MongoDB* database to collect and store the two-dimensional coordinates of 21 finger key points obtained from the finger recognition framework called *MediaPipe*. Furthermore, we organize the experimental scenarios and provide examples of visualizing feature changes through graphs. The outcomes of this research enable the measurement of fine-grained variations in finger motor function, holding the potential for cognitive function assessment and prevention. The remainder of this paper is organized as follows. Section II introduces the preliminaries of the related work and technologies for tapping task. Section III produces a complete description of the proposed method. The implementation and discussion of the proposed Web service for measuring motor function are presented in Section IV, followed by conclusions in Section V.

II. Preliminaries

A. Tapping Task for Measuring Motor Function

One of the methods used to measure motor function quantitatively is the tapping task. The tapping task involves tapping

Authorized licensed use limited to: Kochi University of Technology. Downloaded on March 05,2024 at 01:51:59 UTC from IEEE Xplore. Restrictions apply.

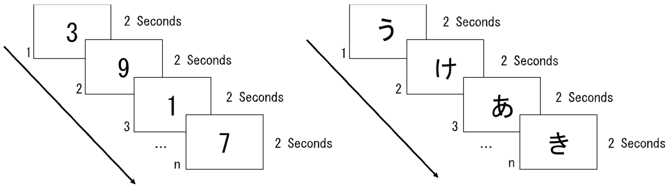

Fig. 1. Examples of presenting number and Japanese Kana in a tapping task.

a pad or a similar surface with one's finger in response to periodic stimuli, often visual. Cognitive and motor functions can be evaluated by assessing how quickly and accurately individuals respond to the presented stimuli [4]. In this study, we propose an overview of a tapping task conducted through a web service. Examples of presenting number and Japanese Kana in a tapping task are illustrated in Figure 1.

Before performing the tapping task, numbers or features are assigned to the fingers of both hands. The web service then presents random numbers or features every two seconds, and participants move the finger corresponding to the presented number to touch a specific area on the screen. The tapping task consists of two variations: the `normal` task, where participants touch the presented number or character immediately, and the `n-back` task, where participants respond to the number or feature presented one step before the current one.

B. Related Work

There are several related works on devices used for measuring tapping tasks. One method explored in research involves using devices equipped with magnetic sensors to measure finger movements during tapping tasks [5]. Magnetic sensors offer the advantage of accurately measuring finger distance data; however, they require specialized equipment, which presents a challenge.

Another measurement method investigated is based on touch panels [6]. When using touch panel devices, there are several challenges to consider. Many touch panels determine input as a binary response (touched or not), and there are limited options for continuously assessing input strength. In addition, these methods do not record movements if a touch is unresponsive. Furthermore, it is worth noting that these previous studies have primarily focused on evaluating only the first and second fingers rather than assessing all five fingers simultaneously.

C. Technical Challenges

We will enumerate the current challenges related to the goal of quantitatively evaluating tapping tasks:

The first challenge involves the difficulty of appropriately recording the measured information. Traditional methods rely on presenting numbers written on paper, for example, as stimuli and visual observation confirm the subject's fingers' movement. While this method does not require specialized tools, it has drawbacks, such as the inability to review results later and reliance on the experimenter's subjective judgment for measurement standards. Additionally, using raw video footage, such as from cameras, is undesirable from a privacy perspective. Methods utilizing touch panels offer the advantage of recording digital information. However, they face the issue of being unable to record finger movements when the subject fails to touch accurately or when a touch goes unregistered.

The second challenge is the requirement for special equipment and procedures. As mentioned earlier, the methods used in the preceding research in Section II-B allow for recording fine-grained data on finger movements but demand dedicated sensors. In conducting tapping tasks, subjects, and experimenters should be able to operate without requiring in-depth knowledge of digital devices. Moreover, the device setup should be straightforward, such as simply turning on the terminal's power.

The third challenge is to facilitate result viewing and assessment. Making tapping task accuracy rates and the distribution of results for each finger easily visualizable would be desirable. It would enable both experimenters and subjects to review the results and make them useful for future reference. For this, proper data recording and processing are necessary.

D. Advantage of IoT Device and Web Server

In this study, we contemplate conducting and evaluating tapping tasks as a web service. There are primarily three advantages to measuring cognitive function using a web service.

The first advantage is leveraging Internet of Things (IoT) devices as input sensors. We use web cameras and touch panels to recognize finger movements. This approach can be implemented inexpensively since no special equipment or facilities are required.

The second advantage is the convenience of operation as it runs within a web browser. We can perform measurements and record data efficiently through simple interactions by conducting stimulus presentation and measurement within a single web application.

The third advantage is managing measurement results using a web server. Storing measurement results on a web server allows for easy comparison of results and facilitates utilizing these results in other services.

E. Collecting Data with Touch Panel

To conduct the assessment of tapping tasks within a web application, touch panel recognition is employed. This involves connecting a tablet-type monitor device equipped with a touch panel to a notebook PC running a web browser, utilizing it as both a display screen and an input device. The recognition of touch panel inputs is implemented in the HTML using the specifications of buttons.

F. Image Recognition Using Pre-Trained Model

In the web application developed in this study, we utilize images of fingers captured by a web camera. To extract feature points from image data and obtain coordinate data, we employ image recognition technology using a pre-trained model. The pre-trained model leverages MediaPipe Hands [7], which is publicly available, to automatically output the 2D XY

Authorized licensed use limited to: Kochi University of Technology. Downloaded on March 05,2024 at 01:51:59 UTC from IEEE Xplore. Restrictions apply.

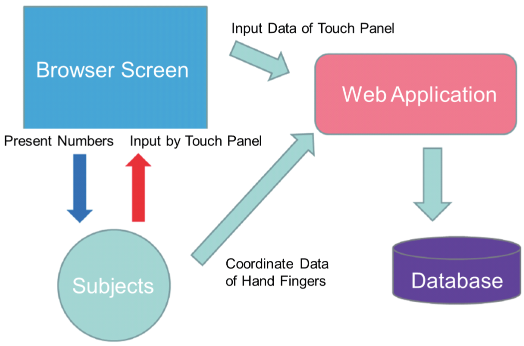

Fig. 2. Overall architecture of proposed method.

coordinates of 21 feature points on fingertip locations from live images captured in real-time.

III. Proposed Method

A. Goal and Approach

The purpose of this study is to simplify the implementation of tapping tasks and enable quantitative measurements. The requirements necessary to achieve this goal are outlined below.

Requirement 1: No Complex Operations Required

Requiring complex operations during the task, not only on the side of the participants but also for the experimenters, can lead to stress during the implementation and decrease experiment speed and measurement accuracy. Hence, it is essential to prominently display only the necessary buttons and input fields, making the operation methods as straightforward as possible.

Requirement 2: Clear Presentation of Results

It is necessary to analyze the obtained results and present them in an easily understandable format. The information to be displayed includes whether the given stimuli match the actions of the participants and the reaction times.

B. Overall Architecture

The overall architecture of the proposed service is illustrated in Figure 2. The proposed service consists of three main components: user registration, experiment execution, and result viewing. It is assumed that only the participants operate the system during the experiment execution phase, while the other components are intended to be operated by experimenters.

C. Module 1: Initial and Experiment Implementation Screen

1) Module 1-1: User Registration Screen: The flowchart for experimenting is depicted in Figure 3. Upon starting the service, the experimenter can choose between conducting the experiment or viewing the results. If they choose to experiment, they input a username and complete the registration process. After registration, the screen transitions automatically. When viewing the results, the experimenter selects the username they registered during the experiment and displays a list of dates and times.

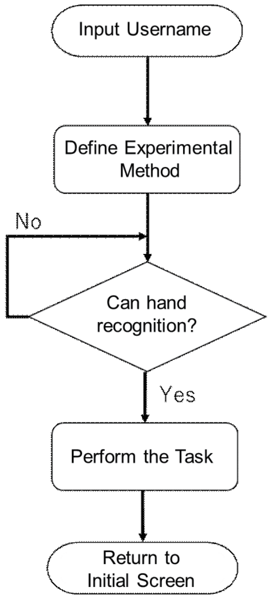

Fig. 3. The flowchart for experimenting.

2) Module 1-2: Experiment Execution Screen: Once the username input is completed on the User Registration Screen, the screen automatically transitions to the Experiment Execution Screen. Next, the experimenter configures information such as the number of times numbers or characters are displayed and whether to perform the `normal` or `n-back` task. Subsequently, adjustments are made to position the subject's hand in front of the webcam to record it. During this process, the experiment can begin once the webcam correctly recognizes the hand. This precaution is taken to avoid starting the experiment without hand recognition, which would result in no data being recorded. Once the hand is correctly recognized, the start button becomes operational. Upon pressing the start button, the tapping task commences. After completing the specified number of displays, the tapping task automatically transitions back to the initial screen.

D. Module 2: Integrating Image Recognition and Touch Panel

1) Module 2-1: Real-Time Replication of Feature Points through Image Recognition: During the experiment, hand image data is captured by a webcam. The captured image data is processed using the `hands.js` library, which extracts only the feature points and records them as coordinate data. While the coordinate data can be overlaid on top of the image data for visualization, it is set to be hidden by default during the experiment.

2) Module 2-2: Algorithm Design for Touch Panel Measurements: The touch panel measurements are performed using ten buttons at the bottom of the experiment execution screen. Participants place their fingers on the buttons and conduct measurements by touching the corresponding button in response to stimuli. The web application determines whether the touched button corresponds correctly to the stimulus and uses this information when displaying results. It is designed to

Authorized licensed use limited to: Kochi University of Technology. Downloaded on March 05,2024 at 01:51:59 UTC from IEEE Xplore. Restrictions apply.

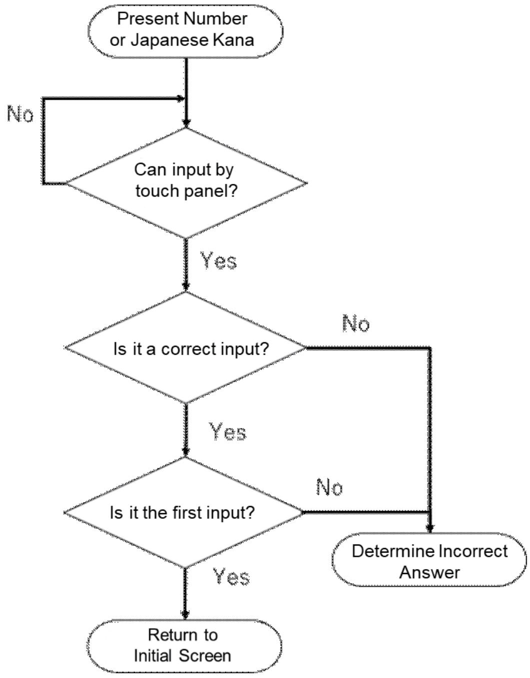

Fig. 4. The flowchart for assessing the input accuracy.

be considered correct only when the given stimulus and button correspond accurately and when the button is not pressed multiple times. If it is correct, the time taken until the button is pressed is recorded as the reaction time. The flowchart of the algorithm for assessing the input accuracy through the touch panel is depicted in Figure 4.

E. Module 3: Viewing Experiment Results

From the initial screen, we can navigate to the screen to view the experiment results. By selecting the username registered during the experiment execution, we can access the results of tasks performed by that user. After selecting a user, we can display the results in JSON format, view a bar graph of accuracy rates, and observe the accuracy trends for each session as a line graph.

F. Building and Linking Databases

The data recorded through measurements is sent to a *MongoDB* database [8]. The transmitted data is divided into hand image and touch panel recognition data, each stored separately. An example format for retrieving data from the database is presented in Table I.

IV. Implementation

A. Technologies Employed

The service proposed in this study is executed in a web browser using `node.js` [9]. We utilized the "Hands" feature within the `MediaPipe` [10] framework to recognize hand images. Furthermore, we incorporated MongoDB to store information in a database for data recording. Table II comprehensively lists the technologies and frameworks utilized in the proposed service.

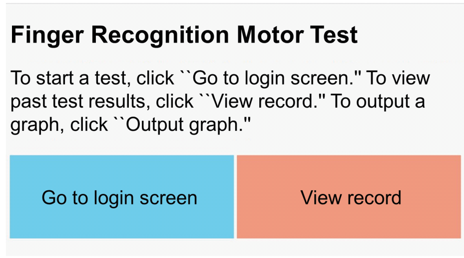

Fig. 5. A screenshot of the service's starting screen.

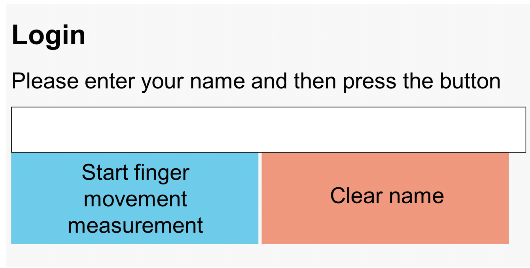

Fig. 6. A screenshot of the username registration screen.

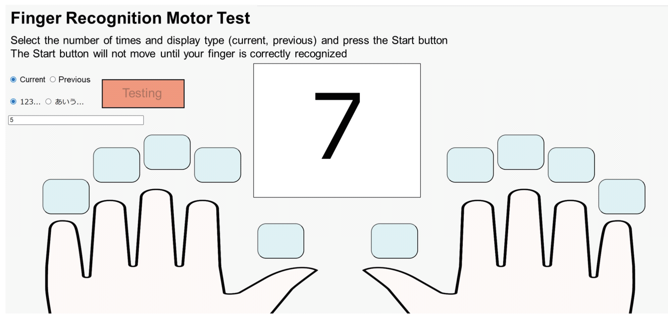

Fig. 7. A screen for experimenting

B. Function Explanation and Workflow During Usage

To begin with, in the initial screen, the experimenter selects whether to conduct an experiment or view results through button input, triggering a branching of scenarios.

Scenario 1: Conducting the Experiment

Figure 5 illustrates a screenshot of the service's starting screen. Initially, the experimenter enters the subject's username and registers it. The registered username is necessary for later reviewing records. Figure 6 displays a screenshot of the username registration screen. Next, the experimenter specifies the format of the experiment and the number of times the stimuli will be presented. Subsequently, after confirming that the subject's hand is correctly recognized, the experimenter presses the Start button to initiate the tapping task. Once the tapping task is completed, the screen automatically transitions back to the initial screen, enabling the next steps. Figure 7 depicts the screen for experimenting, and Figure 8 illustrates the data collection process within the proposed service.

Scenario 2: Viewing Results

In the results viewing section, an example of the user

Authorized licensed use limited to: Kochi University of Technology. Downloaded on March 05,2024 at 01:51:59 UTC from IEEE Xplore. Restrictions apply.

TABLE I
AN EXAMPLE FORMAT FOR RETRIEVING DATA FROM THE DATABASE.

Functions	URL Parameters in `node.js`
Get All UpdateTime	app.get('/userN=:userN/dbN=:dbN/itemN=:itemN/getAllUpdateTime'
Get All Data	app.get('/userN=:userN/dbN=:dbN/itemN=:itemN/getAllData'
Get Specific Data	app.get('/userN=:userN/dbN=:dbN/itemN=:itemN/updateTime=:updateTime/get'

TABLE II
LIST OF TECHNOLOGIES AND FRAMEWORKS USED.

Name	Version
Google Chrome	113.0.5672.127
Node.js	11.0.0
MongoDB	3.5.9
MediaPipe	-

Fig. 8. The data collection process within the proposed Web service.

selection screen in result viewing mode is shown in Figure 9. Users can choose a username and display a data list for each user ID. From this list, they can select the data they wish to view, allowing them to examine input data in JavaScript Object Notation (JSON) format. Figure 10 shows an example JSON file containing touch panel input data, and Figure 11 illustrates an example JSON file containing coordinate data recorded through finger recognition.

Scenario 3: Visualization via Graphs

When viewing results, after selecting a username, users can choose the graph display checkbox to view bar graphs displaying accuracy for each measurement or line graphs showing the accuracy trends. Multiple graphs can be compared side by side by selecting graphs for multiple measurements. Figure 12 displays a screenshot of the graph display screen.

C. Discussion

The service proposed in this study has enabled the rapid execution and recording of experiments. Specifically, input through a touch panel allows real-time correctness assessment during execution, making determining accuracy from the records easy. Additionally, managing user names in the database facilitates the comparison of results from multiple tapping tasks conducted by the same subject. However, the current service has some challenges. Results are output in JSON format, which lacks readability. Furthermore, there is no direct association between the image recognition results and the correctness determination of tapping tasks.

To address the above challenges, several methods can be considered. One approach is to enhance the presentation of results, such as displaying them in easily understandable formats like graphs, which would facilitate comparisons. An-

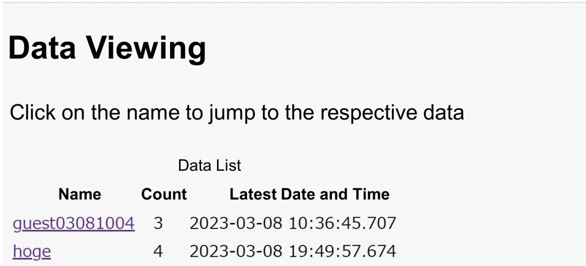

Fig. 9. An example of the user selection screen in result viewing mode.


```
[
    {
        "_id": "6407e450bc32470a14e250be",
        "name": "guest03081004",
        "allData": [
            [
                {
                    "data": 4,
                    "time": "2023-03-08 10:26:27.938",
                    "responseTime": 1063
                }

        ],
        "shownVal": [
            4,
            1,
            8,
            2,
            5
        ],
        "updateTime": "2023-03-08 10:26:40.874",
        "rule": "current",
        "result": [
            1,
            0,
            1,
            1,
            1
        ]
    }
]
```


Fig. 10. An example JSON file containing touch panel input data.

Authorized licensed use limited to: Kochi University of Technology. Downloaded on March 05,2024 at 01:51:59 UTC from IEEE Xplore. Restrictions apply.

```
[
    {
        "_id": "640931b4caba5827085c24f3",
        "name": "i",
        "allData": [
            [
                {
                    "data": [
                        [
                            {
                                "x": 0.6554007530212402,
                                "y": 0.07934701442718506,
                                "z": 4.887745035375701e-7
                            },
                            {
                                "x": 0.6103296279907227,
                                "y": 0.11991216987371445,
                                "z": -0.02835012972354889
                            },
                            {
                                "x": 0.5850882530212402,
                                "y": 0.17868968844413757,
                                "z": -0.04775778576731682
                            },
                            {
                                "x": 0.5738803744316101,
                                "y": 0.24628445506095886,
```


Fig. 11. An example JSON file containing coordinate data recorded.

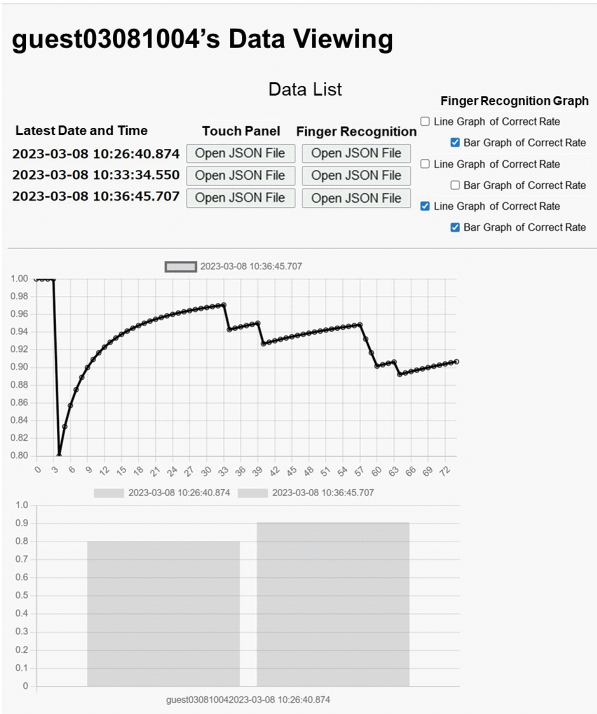

Fig. 12. A screenshot of the graph display screen.

other approach is representing accuracy, making significant results more easily discernible. Additionally, leveraging machine learning techniques may enable the use of hand motion recognition results for correctness determination. By implementing these methods, the service can overcome its current limitations and offer improved functionality for experiment execution and data analysis.

V. Conclusion

In this study, we propose a web service to streamline the execution of tapping tasks. Our proposed service involves recognizing input through a touch panel, recording finger movements through image recognition, and automatically calculating accuracy. It is expected to reduce the workload for experimenters and participants due to its user-friendly features, such as simplified experiment execution and recording. As improvements to our proposed service, we suggest presenting the results of tapping tasks in a more visually accessible format using graphs and image recognition to evaluate tapping task performance. In addition to these enhancements, conducting experiments with participants and incorporating their feedback will refine the service.

Acknowledgment

This research was partially supported by JSPS KAKENHI Grant Numbers JP19H01138, JP20H05706, JP20H04014, JP20K11059, JP22H03699, JP19K02973, and Young Scientists (No.23K17006).

References

[1] A. Kluger, J. G. Gianutsos, J. Golomb, S. H. Ferris, A. E. George, E. Franssen, and B. Reisberg, "Patterns of motor impairment in normal aging, mild cognitive decline, and early alzheimer'disease," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 52, no. 1, pp. P28–P39, 1997.

[2] K. E. Kobayashi-Cuya, R. Sakurai, N. Sakuma, H. Suzuki, M. Yasunaga, S. Ogawa, T. Takebayashi, and Y. Fujiwara, "Hand dexterity, not handgrip strength, is associated with executive function in japanese community-dwelling older adults: a cross-sectional study," *BMC geriatrics*, vol. 18, no. 1, pp. 1–8, 2018.

[3] S. Suzumura, A. Osawa, T. Nagahama, I. Kondo, Y. Sano, and A. Kandori, "Assessment of finger motor skills in individuals with mild cognitive impairment and patients with alzheimer's disease: Relationship between finger-to-thumb tapping and cognitive function," *Japanese Journal of Comprehensive Rehabilitation Science*, vol. 7, pp. 19–28, 2016.

[4] K. Hesseberg, G. G. Tangen, A. H. Pripp, and A. Bergland, "Associations between cognition and hand function in older people diagnosed with mild cognitive impairment or dementia," *Dementia and Geriatric Cognitive Disorders Extra*, vol. 10, no. 3, pp. 195–204, 2021.

[5] J. Sugioka, S. Suzumura, Y. Kawahara, A. Osawa, N. Maeda, M. Ito, T. Nagahama, K. Kuno, K. Shiramoto, S. Kizuka *et al.*, "Assessment of finger movement characteristics in dementia patients using a magnetic sensing finger-tap device," *Japanese Journal of Comprehensive Rehabilitation Science*, vol. 11, pp. 91–97, 2020.

[6] S. Makai-Bölöni, E. Thijssen, E. M. van Brummelen, G. J. Groeneveld, and R. J. Doll, "Touchscreen-based finger tapping: Repeatability and configuration effects on tapping performance," *Plos one*, vol. 16, no. 12, p. e0260783, 2021.

[7] R. K. Veluri, S. R. Sree, A. Vanathi, G. Aparna, and S. P. Vaidya, "Hand gesture mapping using mediapipe algorithm," in *Proceedings of Third International Conference on Communication, Computing and Electronics Systems: ICCCES 2021*. Springer, 2022, pp. 597–614.

[8] K. Saundariya, M. Abirami, K. R. Senthil, D. Prabakaran, B. Srimathi, and G. Nagarajan, "Webapp service for booking handyman using mongodb, express js, react js, node js," in *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*. IEEE, 2021, pp. 180–183.

[9] X. Huang, "Research and application of node. js core technology," in *2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI)*. IEEE, 2020, pp. 1–4.

[10] M. Harris, A. S. Agoes *et al.*, "Applying hand gesture recognition for user guide application using mediapipe," in *2nd International Seminar of Science and Applied Technology (ISSAT 2021)*. Atlantis Press, 2021, pp. 101–108.

Authorized licensed use limited to: Kochi University of Technology. Downloaded on March 05,2024 at 01:51:59 UTC from IEEE Xplore. Restrictions apply.