Estimating Mini Mental State Examination Scores using Game-Specific Performance Values: A Preliminary Study

Hee-Tae Jung¹, Hyunsuk Lee², Kwangwook Kim², Byeongil Kim², Sungji Park³, Taekyeong Ryu³, Yangsoo Kim³, and Sunghoon Ivan Lee¹

Abstract—Individuals with permanent cognitive impairment need to be evaluated and monitored. There exists a number of clinically validated cognitive assessment tools, but they often need to be administered by trained therapists in clinical settings. This serves as a major barrier for frequent, longitudinal monitoring of cognitive function. In this work, we introduce Neuro-World, a collection of innovative 3D mobile games, that allows one to self-administer the assessment of his/her cognitive function. The game performance is analyzed and converted into a clinically-accepted measure of cognitive function, specifically the Mini Mental State Examination (MMSE) score, improving the translational impact of the system in real-world clinical settings. To validate the feasibility of our approach, we collected game-specific performance data from 12 post-stroke patients, which was used to train a supervised machine learning model to estimate the corresponding MMSE score. Our experiment results showed a normalized root mean square error of 5.3% between the actual and estimated MMSE scores. This study enables new clinical and research opportunities for accurate longitudinal assessment of cognitive function via an interactive means of playing mobile games.

I. INTRODUCTION

The number of individuals with permanent cognitive disabilities is increasing in the United States due to the trend of aging society [1]. For these individuals, it is important to evaluate and longitudinally track their cognitive functional/impairment level [2], which is often done using clinically validated assessment tools, such as the Mini Mental State Examination (MMSE) [3] and the Montreal Cognitive Assessment (MoCA) [4]. Unfortunately, these tests need to be performed under the supervision of trained staff in clinical settings. This may negatively affect the productivity of the staff, especially when clinical resources are limited and patients need to be frequently assessed [5].

Addressing the constraint of requiring the presence of trained clinical staff during in-person assessment sessions, various approaches have been investigated, including remotely administering the conventional assessment tools [6], [7] and self-administered computer-based or mobile-device assessments [8], [9]. Remote cognitive assessment tools have great potential to enable patient-centered care by frequently monitoring individual-level outcomes over time in patients’ naturalistic environments [9]. Brandt et al. devised a new questionnaire that can be administered by trained clinicians over the phone, so that follow-up assessments can be made more easily [6]. Similarly, recent work by Absolahi et al. demonstrated the feasibility of administrating the MoCA via video conferencing [7]. Although these approaches enable frequent, remote assessment of patients, they still require the presence of trained clinicians and have minimal impact on the optimal utilization of clinical resources. There have been attempts to enable self-administered cognitive assessments using computers or smart mobile devices [9], but they stop short at reporting software-specific assessment scores rather than the widely accepted assessment scores, such as MMSE or MoCA. This poses a significant translational limitation, and it is necessary for clinicians to undergo additional training to interpret their results.

In this study, we introduce Neuro-World, that we developed specifically to challenge and measure confounding cognitive function in patients. We further propose a supervised machine learning approach that can translate the game-specific performance features, that were extracted from patients’ engagement with Neuro-World, to the MMSE score directly. The performance of the proposed system was evaluated based on the data collected from a total of 12 post-stroke patients with cognitive impairment who visited the clinic twice (i.e., at baseline and follow-up after 3 months). The results suggest that Neuro-World and the proposed approach can be self-administered by patients and accurately estimate MMSE scores, opening up new possibilities to assess cognitive function and impairment.
TABLE I
THE DESCRIPTION OF THE 6 GAMES FROM THE NEURO-WORLD USED IN THIS STUDY

<table>
<thead>
<tr>
<th>Game #</th>
<th>Title</th>
<th>Presentation &amp; Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Perception Memory</td>
<td>Animals come into the screen and an animal leaves the screen in one of the four directions (up, down, left, right). The subject is asked to remember and choose the correct direction that the animal left.</td>
</tr>
<tr>
<td>2</td>
<td>Object Memory</td>
<td>Animals come into the screen one by one and a subset of them leaves the screen one by one. The subject is asked to remember all the animals which left the screen and choose them in the presented candidates.</td>
</tr>
<tr>
<td>3</td>
<td>Sequential Memory</td>
<td>Animals come into the screen in a specific order. The subject is asked to choose the correct order that the animals entered the screen.</td>
</tr>
<tr>
<td>4</td>
<td>Selective Attention</td>
<td>A sequence of shapes is presented on the screen in a certain pattern. The subject is asked to choose the correct shape that should come after the displayed sequence among the candidates.</td>
</tr>
<tr>
<td>5</td>
<td>Vigilance Attention</td>
<td>A set of animals are displayed on the screen. The subject is asked to count the number of animals that have the same visual characteristics, that is described in a separate image.</td>
</tr>
<tr>
<td>6</td>
<td>Visual Investigation</td>
<td>A set of animals are presented on the screen. The subject is asked to find the same animal that is presented in a separate image among the presented animals.</td>
</tr>
</tbody>
</table>

II. MMSE & NEURO-WORLD

MMSE consists of a set of questions that evaluate different aspects of cognitive function, which are categorized as orientation, registration, attention & calculation, recall and language [3]. Although these categories of MMSE may seem to measure distinct aspects of human cognition, their performances are the combining results of basic cognitive processes, such as memory, attention and visual perception [10], [11]. This understanding served as our design principle while developing Neuro-World.

Neuro-World consists of 6 mobile games, each of which was designed to challenge memory, attention and visual perception in patients. The descriptions of the games are summarized in Table I. Patients start each game at the lowest difficulty and proceed to the next level where the difficulty is adjusted in various ways based off of psychological understanding [12]: increasing the number of animals, their speed entering and exiting the screen, the number of answers in the multiple choices, etc. Each game will stop when patients fail to complete a particular stage multiple times in a row. Hence, the number of stages completed and the time that it took for patients to clear all stages will be subject to the level of patients’ cognitive function.

III. DATA COLLECTION

The study was approved by the Institutional Review Board. Inclusion criteria for this particular study stipulated the recruitment of post-stroke survivors with mild cognitive impairment in their chronic stage (2 years or longer since their last onset) from Heeyeon Rehabilitation Hospital, South Korea. Subjects were first evaluated for their cognitive impairment level by using the MMSE administered by a trained therapist. Subjects who scored 19 or higher on the MMSE were included in the study to ensure that they could understand instructions and play the game themselves. To ensure that the measured game-specific performance values are subject to cognitive function, patients with significant visual or motor impairments were excluded from the study. As a result, a total of 12 post-stroke subjects (73.6 ± 7.3 years old) participated in the study. All subjects were Asian. 91.67% were female and had ischemic stroke. One patient had Alzheimer’s in addition to stroke. No subjects had experience with Neuro-World prior to the experiment. Table II summarizes the information of the recruited subjects.

All the experiments and examinations were conducted at Heeyeon Rehabilitation Hospital only when patients felt comfortable. Subjects were asked to take comfortable sitting positions in places of their choice, e.g., their rooms, corridors, and any public space in the hospital. Neuro-World was installed on an 12.2-inch tablet device (Galaxy Note Pro, Samsung). Researchers explained the graphical interface of the application and asked the subjects play a single stage for each game to verify if they had any difficulty understanding the graphical and/or textual information provided on the screen. Then, subjects were instructed to play all 6 games from the first (easiest) stage and continue until they reached the stage of their maximum cognitive capacity. In general, patients spent approximately 5 minutes engaging with each game, totalling 30 minutes for the entire Neuro-World. After the completion, Neuro-World stored game-specific values reflecting subjects’ cognitive function in the tablet device, which will be explained in Section IV. After the initial data
collection, the subjects were allowed to play the games for 30 minutes a day, 2 days a week, for 3 months. This was done to investigate the impact of extensive exposure to the game and its learning effect on the MMSE estimation. Our hypothesis was that the game-specific values do not overestimate the cognitive level due to patients getting used to the games’ strategies. Afterwards, another round of data collection was done using the same procedure, yielding a total of 24 game plays and MMSE evaluations in the data set.

IV. DATA ANALYSIS

A. Overview

We aim to produce estimation models that correlate game-specific variables, that were extracted from patients’ self-administration of Neuro-World, directly to the MMSE score through a supervised machine learning approach.

B. Game-Specific Features

Game-specific variables that we hypothesized to be relevant to patients’ cognitive function were extracted from the games played by our subjects. The following 12 variables were extracted from each of the 6 games, yielding a total of 72 features.

1) The stage of the highest difficulty level that a subject reached ($x_1$).
2) The shortest time, longest time, and mean time taken for a subject to clear any of the stages ($x_2, x_4$).
3) The shortest time, longest time, and mean time taken for a subject to start selecting answers for any of the stages ($x_5, x_7$).
4) The total number of stages that a subject tries, clears, and fails ($x_8, x_{10}$).
5) The averaged score of all the attempted stages that describes the overall performance in each stage ($x_{11}$).
6) The overall score of a subject’s performance for the entire game ($x_{12}$).

Detailed measures used to compute $x_{11:12}$ were proprietary and not available at the time of this project. It is difficult to isolate individual assessment of memory, attention, visual perception in each individual feature, and our choice is to capture their confounding effects in a variety of features.

C. Classification and Evaluation

First, the correlation-based feature selection algorithm [13] was applied to identify a subset of game-specific variables that were highly correlated to the MMSE score and to avoid overfitting. Then, a multi-class classification model was trained using the Random Forests algorithm to estimate the corresponding overall MMSE scores. Random Forests is an ensemble learning method that constructs multiple weakly-correlated decision trees based on randomized features and bootstrap samples [14]. The algorithm produces the predicted class label by averaging the decisions of the trees. By averaging many noisy and approximately unbiased models, the algorithm can reduce the variance of the prediction function and thus minimize the chances of overfitting [14].

V. RESULTS

The overall estimation performance was assessed using the leave-one-subject-out cross validation (LOSO-CV) technique. Hence, 12 iterations of cross validation were performed, in which the data obtained from 11 patients (collected at both baseline and follow-up) were used to train an estimation model and the data belonging to the left-out patient (again, the baseline and follow-up data) were used to evaluate the trained model. The root mean square error (RMSE) and normalized RMSE (NRMSE) were used to evaluate the estimation performance. All the analyses were performed in the MATLAB R2016b environment with the matlab2weka interface to Weka 3.6.12.

The RMSE between the actual MMSE scores administered by trained clinicians and the estimated scores based on the game-specific variables was 0.58. The normalized RMSE, computed by dividing the RMSE by the value range of the actual MMSE score (i.e., $30 - 20 + 1 = 11$), was 5.3%, showing that the proposed system could yield an accurate evaluation of MMSE. Fig. 2 shows a scatter plot between the estimated vs. actual MMSE scores. We had 8 instances that were incorrectly classified out of 24 instances. However, all of the incorrectly classified MMSEs deviated from the actual score only by 1 point. It is noteworthy that 4 of these 8 incorrectly classified instances were the only data points within their classes (i.e., our data set had only 1 instance that scored MMSE of 21, 25, 28, and 30, respectively), but our algorithm was able to reliably estimate these scores to the nearest possible MMSE scores. This indicates that our game-specific variables can capture the varying degrees of cognitive functions measured by the MMSE.
Towards that end, we developed a collection of 3D mobile games, Neuro-World, that can translate game performance into the MMSE score using a machine learning algorithm. The presented results obtained from 12 post-stroke individuals with mild cognitive impairment show that our approach can accurately estimate the MMSE score (i.e., NRMSE of 5.3%). The proposed system has potential to allow follow-up assessments to be performed remotely in a more convenient and user-engaging manner.

ACKNOWLEDGMENT

The authors would like to thank the staff members from Heeeyeon Rehabilitation Hospital, S. Korea, and the engineers from Woorisof, S. Korea, for their support throughout the study. The authors appreciate Brandon Oubre and Kallie Whitenour for their help in proofreading the manuscript.

REFERENCES


**TABLE III**

The top six most frequently selected features that were used to estimate the MMSE score throughout the 12 LOSO-CV iterations.

<table>
<thead>
<tr>
<th>Game #: Title Feature</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>5: Vigilance Attention Total number of cleared stages</td>
<td>12</td>
</tr>
<tr>
<td>6: Visual Perception Scaled value of summarized score</td>
<td>10</td>
</tr>
<tr>
<td>1: Perception Memory Scaled value of summarized score</td>
<td>10</td>
</tr>
<tr>
<td>6: Visual Perception Longest time taken to start selecting answers</td>
<td>9</td>
</tr>
<tr>
<td>4: Selective Attention Scaled value of summarized score</td>
<td>5</td>
</tr>
<tr>
<td>5: Vigilance Attention Shortest time taken to clear a stage</td>
<td>2</td>
</tr>
</tbody>
</table>

The mean and standard deviation of the number of features used to train a classifier was 6.3±1.7 among the 72 features considered in this work. Table III presents the top 6 most frequently selected features throughout the 12 LOSO-CV iterations. It is interesting to note that the features from games 1, 5, and 6 were used to train classifiers in 11 out of 12 iterations. The features from game 4 were selected 6 times.

On the other hand, features from games 2 and 3 were used only twice out of 12 iterations. This implies that the games 1, 5, and 6 may contain sufficient information to yield an accurate estimate of the MMSE score. These results provide a potential opportunity to utilize only a subset of games to estimate the MMSE score in reduced time.

VI. DISCUSSION

Our results indicate that it will be possible for patients with mild cognitive impairment to play Neuro-World outside the clinical setting and that the assessment of MMSE can be automated. This overcomes the traditional constraint of requiring the supervision of trained clinicians, which can enable a more optimized utilization of clinical resources.

Neuro-World provides an opportunity to assess cognitive function via a series of interactive games, which we assume to make the process more engaging and fun, and thus, improve patient adherence to longitudinal and frequent assessment. This, in turn, may make it possible to detect cognitive disturbance among potential patients early on, highlighting the translational impact of our system in real-world settings.

When compared to paper-based tests and their digitized versions, it would be easier to randomize the visual or movement patterns of the presented stimuli in game-based tests. This may actually help prevent learning effect and achieve reliability in frequent re-tests.

VII. CONCLUSION

In this work, we introduced a novel means to accurately estimate clinically validated cognitive assessment scores that can be self-administered without the supervision of trained clinicians. Towards that end, we developed a collection of