Collecting Digital Biomarkers on Cognitive Health Through Computer Vision and Gameplay: an Image Processing Toolkit for Card Games

Karsten Gielis
e-Media Research Lab – KU Leuven
Department of Computer Science
Leuven, Belgium
karsten.gielis@kuleuven.be

Joren Kennes
KU Leuven
Fac. Engineering Technology
Leuven, Belgium
joren.kennes@student.kuleuven.be

Christophe De Dobbeleer
KU Leuven
Fac. Engineering Technology
Leuven, Belgium
christophe.dedobbeleer@student.kuleuven.be

Steven Puttemans
EAVISE – KU Leuven
Department of Electrical Engineering
Sint-Katelijne-Waver, Belgium
steven.puttemans@kuleuven.be

Vero Vanden Abeele
e-Media Research Lab – KU Leuven
Department of Computer Science
Leuven, Belgium
vero.vandenabeele@kuleuven.be

***Abstract*— Individuals suffering from cognitive impairment require early diagnosis and frequent follow-up. However, as the healthcare system is overburdened, contact moments with patients are sparse. Digital biomarkers are a form of non-intrusive continuous measurements of cognitive health, found in daily activities. They can help in the diagnosis and follow-up of cognitive impairments. Playing videogames is an example of these daily activities which requires cognitive effort, hence carrying cognitive information. However, extracting these digital biomarkers from commercial games tends to be challenging. To this end, we explored capturing digital biomarkers from existing card games through computer vision. A toolkit was designed for the standard Microsoft Solitaire Collection, allowing for unobtrusively measuring digital biomarkers. Results show that this technique allows for real-time processing of cognitive digital biomarkers. Digital biomarkers were captured from 44 participants in three age groups. An initial data exploration supports the promise of these digital biomarkers as bearers of cognitive information. Differences were seen amongst the age groups caused by age-related cognitive changes. These results suggest that digital biomarkers in commercial games can be used for unobtrusive long-term cognitive monitoring with minimal burden on patient and physician, potentially leading to a more complete, clear cognitive profile.**

Keywords*— *Digital Biomarkers, Cognitive Health, Game Analysis, Card Games, Image Processing

I. Introduction

Worldwide millions of people suffer from cognitive disorders such as depression (322 million), anxiety (264 million), or dementia (50 million)[1], [2]. Even milder cognitive impairments such as Mild Cognitive Impairment (MCI) can hamper several cognitive functions such as attention, executive functioning or social cognition. Depression alone accounts for 4.3% of the Global Burden of Disease, making it the largest cause of disability worldwide [3]. The cost of dementia, not including the emotional stress on families, was estimated at 818 billion dollars in 2015 [1]. Anxiety and depression combined account for a global cost of 1.15 trillion dollars per year [4].

Hence, early diagnosis and frequent follow-up of mental health problems is crucial to managing the disease, allowing for timely treatment and disease progression mitigation. It ensures finding the best sources of support and making informed decisions about the future, even if the disorder is untreatable [5]–[9]. For some cognitive ailments, especially dementia, diagnosis is often non-existent or made in a later stage of the disease. A study in 2015 reported that 58% of all dementia cases in the USA go undiagnosed [10]. This is why, in 2017, the World Health Organization (WHO) endorsed the Global action plan on the public health response to dementia focusing amongst others on diagnosing cognitive impairments in an earlier stage.

A considerable part of diagnosis and follow-up in traditional medicine involves the use of biomarkers. Biomarkers are defined by the WHO as "any substance, structure, or process that can be measured in the body or its products, and that influences or predicts the incidence of outcome or disease" [11]. They are an objective way to indicate biological and pathogenic processes or responses to therapeutic interventions, utilized in the fields of

978-1-5386-9138-0/19/$31.00 ©2019 IEEE

disease prediction, diagnosis, and prognosis [12]. A well-known biomarker used in medical diagnosis is e.g., the presence of Amyloid beta in cerebral spinal fluid for Alzheimer's Disease [12],[13]. Asides from biomarkers to aid in the screening, diagnosis, and follow-up of mental health illness, which are often expensive and invasive, neuropsychological tests are common practice. Well-known cognitive tests to aid in the screening for cognitive impairments are e.g., the Mini-Mental State Examination (MMSE) or the Montreal Cognitive Assessment (MoCA) are used to screen for dementia [14], [15]. While these tests are less invasive and less expensive than capturing biomarkers, they are also characterized by lower specificity and sensitivity [16]–[22]. That is why there is an increasing interest in digital biomarkers. Digital biomarkers are defined as *"objective, quantifiable physiological and behavioral data that are collected and measured by means of digital devices such as portables, wearables, implants or digestibles"* [23].

Digital biomarkers have the possibility to give deeper insight to specialists and patients, providing a source of data from text interactions, home data, GPS location, but also *games*. Mandryk and Birk point out that a variety of activity traces can be gathered from in-the-wild gameplay of COTS games and considered as *digital biomarkers of cognitive health* [24]. The contribution of this paper is to explore gameplay as an additional medium to capture digital biomarkers for mental health. In particular, this paper presents a new method of defining digital biomarkers in games and a toolkit for collecting digital biomarkers of cognitive impairment through gameplay. The toolkit and method are developed to work with existing card games, i.e. the Microsoft Solitaire Suite, that people already play and enjoy. Finally, an exploratory data analysis is shown to demonstrate the feasibility of the toolkit.

II. Background

In this section, we first explore the potential digital biomarkers for mental health. Next, we present games as a viable source for capturing these digital biomarkers. Next, we explore challenges when games are used that are developed by research labs. Finally, we present the opportunities of using commercial-of-the-shelves games as a source of digital biomarkers.

A. Digital Biomarkers

Today, in the cross-domain of computer science, engineering, biomedicine, regulatory science, and informatics, interest is growing in the digital counterparts of biomarkers. The surge in interest has sparked the founding of a journal by Karger in September 2017 [23] dedicated to this topic only. Compared to classical pen-and-paper tests, the use of digital biomarkers has shown advantages such as reduced cost, unobtrusive measurement and the possibility of continuous data gathering. In contrast to episodic measurements of classical biomarker and pen-and-paper tests, which are often taken biannually or yearly, digital biomarkers can be captured on a daily basis. This makes the findings more robust to patients having a momentary lapse, feeling stressed, examined, or being tired because of a bad nights rest. It has also been shown that pen-and-paper tests are vulnerable to practice effects due to the fixed course of the tests [25]–[27]. More reliable and comprehensive cognitive test batteries do exist, but these lengthy tests have to be administrated by a trained health professional and require the often frail participant to go to a specialized institution. In these specialized institutions, often one specialist is assigned to follow-up the disease progression as intra-rater reliability is proven to be better than inter-rater reliability [28]. However, even these specialists suffer from bias, reducing the overall precision of these follow-up tests [29]. Moreover, as these specialists are overburdened, contact moments can be sparse. As shown by the Digital Biomarkers Department of Roche in Fig. 1, sparse contact moments can cause neuropsychologist to miss the bigger picture as patients tend to recall symptoms for smaller time periods than the follow-up period actually spans [30]. For people suffering from mental degeneration, remembering and evaluating the severity of symptoms can prove even more difficult.

Digital biomarkers possibly solve many of these problems. They may increase the ecological validity by increasing the temporal and spatial resolution of the captured behavior during activities of daily life [31]. Secondly, they may unlock previous unobtainable sources of behavioral, social, environmental, and physiological data [32] with minimal effort required from physician and patient. Finally, as these digital biomarkers are captured by digital devices, they are less prone to human bias and less susceptible to the white-coat effect[33].

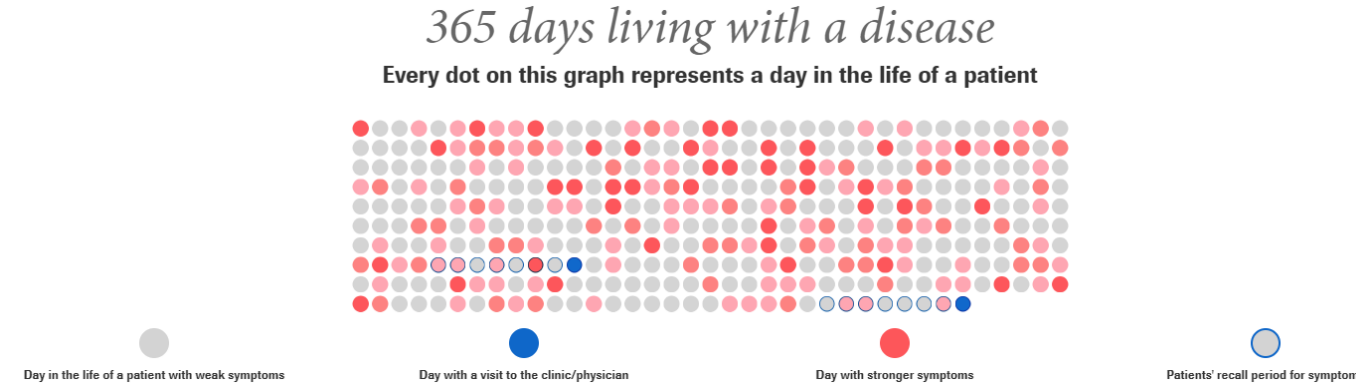

Fig. 1. Limited contact moments can cause physicians to miss the bigger picture [18]

The advantages of digital biomarkers have sparked several initiatives on diverse platforms using various sensors. Redfield et al. used accelerometers of smartphones to measure gait, finger tapping, voice, and balance as a measure for Parkinson disease [34]. Saeb et al. found strong evidence that mobile phone sensor data such as GPS and phone usage correlates to depression [35]. Faurholt-Jepsen et al. and Beiwinkel et al. did research on biomarkers and bipolarity, correlating smartphone information such as calls, text messages, and GPS with mental disorders [36][37]. In a study using the Beiwe app, schizophrenia relapses were correlated with anomalies in patient behavior prior to relapse [38]. Hagler et al. created an in-home monitoring system to assess gait, a predictor of cognitive decline [39]. While in most cases the sample sizes were too small to draw conclusions about the general population, these experiments show promise of continuous monitoring of at-risk populations with minimal effort required from the user.

B. Digital Biomarkers and Games

Games are a natural source of information on behavior, cognitive performance, motor performance, social behavior and affect, for people of all ages [24]. According to Suits, games are interactive systems, where players present *"a voluntary attempt to overcome unnecessary obstacles"*[40] Through playful design, and intuitive rules, players are motivated to push their own boundaries [41]. In certain cases, it can even cause a state of *flow* [42], a gratifying state where players lose track of time and place because the challenging game activity necessitates all of their attention and skills. This flow experience is found enjoying and makes games *autotelic*, players are intrinsically motivated to play [40]. As research has shown that decrease in motivation has an effect on participant performance of classical pen-and-paper cognitive tests[43], the potential of games to maximize motivation makes them a suitable medium for capturing digital cognitive biomarkers.

There are already many examples of gamified collect digital biomarkers which may be indicative for cognitive performance. Episodix, a gamificated California Verbal Learning Test, manages to classify individuals into three categories: healthy, mild cognitively impaired, and Alzheimer Dementia [44]. A mirror game designed by Słowiński et al. captures digital diagnostic biomarkers in the form of non-verbal synchrony and neuromotor functions. Utilizing statistical learning techniques, they could discern users suffering from schizophrenia from their healthy counterparts, with an accuracy of 93% and a specificity of 100% [45]. Neuro-World, a collection of 3D mobile games by Jung et al. [46], estimates Mini-Mental State Examination scores from gameplay metrics such as score, time to clear a stage, and the number of cleared levels. Their games test perception, object memory, sequential memory, selective attention, vigilance attention, and visual investigation. Leduc-McNiven et al. developed WarCAT and Lock Picking [47]. WarCAT, a card game based on War, measures recognition and recall while Lock Picking measures problem-solving skills by letting the user search for an optimal score. Smartkuber, the augmented reality game for cognitive screening made by Boletsis and McCallum [48] uses five minigames to screen for cognitive impairments, revealing significant correlations and comparable validity to the Montreal Cognitive Assessment, a popular neuropsychological screening test for Mild Cognitive Impairment.

C. Challenges with serious games

These games fall under the category *serious games,* defined as "Games that a serious goal, rather than entertainment, enjoyment or fun, as their primary purpose"[49]. In this case, their primary purpose is to provide information on cognitive performance. Unfortunately, research shows that these tailor-made *serious games* suffer from disadvantages. First, these games made in research labs often miss the funding and development time of commercial games. Developing a serious game that can compare to commercial games in quality, is often out of reach due to differences in manpower, budget, and expertise. As research cycles differ from game release schemas, it is likely that the game will be outdated by the time the game is programmed, funding is gathered, and medical ethical clearance is approved [50]. Maintaining the game and shipping updates also prove difficult as this is not the main goal of research labs.

Secondly, despite the efforts to make them as enjoyable as possible, research has shown that custom-made games for cognitive training still fall short in engagement and suffer from attrition in longitudinal studies. The repetitiousness of many gamified assessments and training can lead to participant disengagement, possibly impacting the data quality [51] [52] Furthermore, it has been reported that it is the affectionate bond with the experimenter and not the cognitive training per se that motivated participants to continue [53]. This suggests that there is a mismatch between the serious games being developed for cognitive functioning and the games people effectively enjoy playing. It may be that serious games, while valid with respect to the mental health purpose, perhaps do not provide 'meaningful play'[54].

D. Opportunities of Commercial-Of-The-Shelves games

The playing of commercial games is weaved into the fabric of everyday lives; such games are part of the socio-cultural environment[54]. As mentioned above, the power of digital biomarkers lies in its frequent, longitudinal measurement, stressing the importance of the autotelic nature of games. Boot et al. discovered during post-intervention surveys that the games of the control condition, such as word and puzzle games, were found more enjoyable than those of the gamified test group [55]. This enjoyment of the game led to higher motivation to adhere to the cognitive training, indicating that commercial games may be a better fit for capturing digital biomarkers. This may make COTS games a more valid, suitable medium to gather digital biomarkers for cognitive performance [56]. However, the downside of using these commercial, off-the-shelf (COTS) games is that they gameplay is less 'controlled', they may demand more complex and variable actions from players, simultaneously addressing multiple cognitive functions from the players. This interplay of different cognitive functions may introduce undesired non-therapeutic effects or add uncertainties in screening and in detecting impairments in cognitive functioning [57].

We are not the first to promote the use of commercial, off-the-shelf (COTS) games for assessment of cognitive performance. Jimison et al. showed that there is a correlation between Mild Cognitive Impairment and performance in the game FreeCell. They found that higher variability in scores and more sensitivity to game difficulty are indicative of cognitive impairment[27]. Thompson et al. explored the relation between common puzzle games and standard neuropsychological tests and found that performance on these smartphone-based games is indicative of cognitive ability across several cognitive domains[56]. Furthermore, working memory was correlated to sudoku performance by Grabbe, showing the potential of this popular game for measuring cognitive performance [58].

However, as aforementioned, capturing digital biomarkers in COTS games can be troublesome, as altering the code of the game is impossible without the permission of the game developer. An alternative is recording and annotating gameplay manually, reviewing hours of gameplay and manually tagging digital biomarkers. Yet, manually annotating game data reintroduces the limitations of the aforementioned classic tests. It reintroduces human error, as manually timing of events is more inaccurate and inconsistent. Secondly, manually annotating is a time-consuming and tedious task. Finally, it limits the number of metrics captured, it refrains from capturing certain digital biomarkers. More fine-grained biomarkers such as speed or certainness of execution are not measurable from manually annotating gameplay alone. Moreover, if the researchers want to explore previously uncaptured biomarkers, the annotating process starts anew.

This paper explores a different method for capturing digital biomarkers, namely computer vision and more specifically image processing algorithms. By utilizing machine learning and image processing, gameplay can be analyzed in real-time and digital biomarkers can be extracted and processed in an efficient manner. In order to explore the viability of image processing to capture digital biomarkers on a COTS game, a multithreaded C++ desktop application was developed that utilizes the Open Source Computer Vision Library (OpenCV) [59]. It is built as a generic toolkit to capture and analyze card game play data. It acts as a silent watcher which unobtrusively captures, processes, and analyses gameplay from the standard Microsoft 10 Solitaire Collection. Currently, the code of the game rules is implemented for Klondike Solitaire and FreeCell versions. In the next sections, we will demonstrate how the toolkit operates by using *FreeCell* as an example.

III. Toolkit Concepts

This section contains information necessary to grasp the mechanisms behind the toolkit. First, the rules and board space of FreeCell are illustrated. Secondly, the method of defining digital biomarkers is explained. Finally, the global concept of capturing digital biomarkers in gameplay is described.

A. The FreeCell Board Space

FreeCell is a well-known and popular Solitaire variant. It is played with all 52 cards in a deck, which are all dealt face-up, at the beginning of the game. This transparency of the board makes that almost all FreeCell games are solvable. Of the original 32000 different starts of the FreeCell game (the Microsoft 32K variant), only one is deemed impossible to solve, making approximately 99.99% of all FreeCell deals solvable [60].

As seen in Fig. 2, the playing board consists of three parts. The large section at the bottom is called the *build stack*, where all fifty-two cards reside at the start of the game, divided over eight stacks. The part at the top left of the board is called the *storage stack*. Here, a card can be temporarily stored during the game. The last section of the board, at the top right, is called the *suit stack*. The goal of the game is to move all the cards here. Playing cards comprise four suits: clubs, diamonds, hearts, and spades. On the corresponding suit stack, the cards need to be placed per suit in ascending order: starting with the ace, then two, three etc., ending with the king. When all the cards are placed on the suit stack, the game is won.

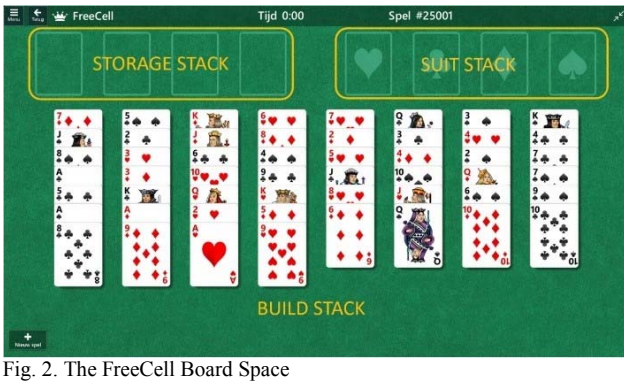

Fig. 2. The FreeCell Board Space

To accomplish this goal, some rules have to be followed. It is allowed to move cards from one *build stack* to another if 1) its rank is one lower than the current top card of the pile and 2) of the opposite color. For example, a nine of (red) hearts can be placed on a ten of (black) clubs. The general rule is that only one card is allowed to move at once. However, cards moved on top of each other with alternating colors and descending rank are allowed to move to a new location, given that there are enough free spots on the *build stack* and/or *storage stack*. The maximum number of cards that are allowed to be moved in one single move can be calculated with the following equation: $\#\ \text{movable cards} = (1 + \#\ \text{free spots}) \times 2^{(\#\ \text{free columns})}$.

B. From Game to Digital Biomarker

As aforementioned, measuring digital biomarkers from COTS games can be troublesome since these games simultaneously require multiple cognitive functions, in contrast to gamified tests which are custom made to capture a specific cognitive function. Therefore, it is imperative to outline the methods that were used to explore, extract and define specific digital biomarkers from gameplay. In order to translate gameplay into digital biomarkers, we applied a methodical approach existing of three phases. We first started by creating an exhaustive list of game events. In the second phase, we converted them into player mistakes. In the third phase, we quantified these mistakes to transform them into possible digital biomarkers.

For the first phase, two researchers in the field of human-computer interaction (KG and VV, co-authors of this paper) and two master students (JK and CD, also co-authors of this paper) created a list of all possible game events for the game FreeCell. The literature on the topic of FreeCell and its rules was gathered, studied, and processed [61]–[66]. This literature ranged from optimal solvers, to previous cognitive studies, to hardness analysis. This gave insight into the common pitfalls, optimal solving strategies, and cognitive studies previously done on the subject. Next, to this literature study, the game was played in several sessions. In a series of iterations, a list of game events was drafted and refined until no more game events were found. Through this processed literature, in combination with the information gathered through the extensive gameplay, a thorough, comprehensive list of game events was generated. These game events consisted among others of game outcomes (e.g. game won or lost), player moves (e.g. storing a card in the *storage*), and incorrect player moves (e.g. placing a card on another card with the same color on *the build stack*).

In the second phase, to reduce this list and prevent duplicate records for the same event (e.g. positively and negatively phrased game outcomes), all game events were converted as player actions that may be indicative of cognitive impairment. For example: 'User makes a correct move' was translated into the player mistake 'User makes an incorrect move'. Next, player actions were further specified. For example 'User makes an incorrect move' was further detailed, into 'User makes a rank error' and 'User makes a suit error'. This resulted in a set of 16 possible player actions indicative of cognitive performance. Next, this set of player actions was reviewed again, and only those actions that can be captured unambiguously via playing behavior were retained. Therefore, player actions which required insight into the current mindset of the player were not captured. Additionally, only player actions which are unquestionably erroneous remained. For example, "Player stores a card with no clear advantage." was omitted as well.

Finally, in the third phase of our systematic approach, these remaining player actions were quantified. In other words, for each player action, the measurable element was determined as well as the type and range of (i.e. the game outcome that is measurable on a quantitative scale). These final elements are considered as potential digital biomarkers as they can be unambiguously captured and are potentially influenced by cognitive status. For FreeCell, 10 digital biomarkers were defined. Next to these biomarkers, metadata concerning the games and moves are captured. Table 1 shows all 10 digital biomarkers and metadata as captured in FreeCell.

All digital biomarkers are designed to measure at the lowest level as possible. In this manner, they can become the building blocks of more complex composite digital biomarkers. For example, digital biomarkers such as 'Think time before making an erroneous move' or 'longest error streak', can always be extracted as a combination of these original digital biomarkers. Furthermore, metadata of the x- and y- coordinates can be used to calculate the speed of moves. By capturing this information at the lowest level, there are few limitations on the number of derivates or combinations of digital biomarkers.

C. *Efficiently Capturing Digital Biomarkers*

To efficiently capture digital biomarkers, the toolkit should not process images when the user is not interacting with the game (e.g. thinking of the next move). Therefore, event-driven interrupts are programmed to ensure optimal performance. These events are triggered by the user and consist of a combination of keyboard, controller or mouse input. As FreeCell is solely played with the left mouse button, the event-driven interrupts consist of left-clicking, double-clicking, and dragging.

TABLE 1. DIGITAL BIOMARKERS AND METADATA CAPTURED IN FREECELL

Digital Biomarker	Explanation	Value
Suit Error (SE)	This error is prompted when a card is placed on another card with incompatible suits.	total
Rank Error (RE)	This error is prompted when a card is placed on another card with incompatible ranks.	total
Moved Too Many Cards Error (MMCE)	This error is prompted when a card or a group of cards is moved when there is not enough room to execute said move.	total
Unmovable Card Error (UCE)	This error is prompted when the user tries to move a card which is unmovable (i.e. there are still cards above the card that need to be moved before the original card can be moved).	total
Think Time (TT)	Think Time is defined as the time between the last card placed and the first card touched to make a new move.	ms
Move Time (MT)	This is the time necessary for a user to move a card from one place to the other.	ms
Game Result (GR)	The outcome of the game, whether the user was able to place all cards on the *four suit stacks* and won the game.	WON/ LOST
End of Game (EoG)	Whether the user gave up or the game indicated that there were no more moves.	YES/NO
Number of undo's (NU)	The number of undo's requested by the user.	total
Number of hints (NH)	The number of hints requested by the user.	total
Move Details	Metadata of each move is stored such as x- and y-coordinates, the selected card, source location, destination location and the number of cards moved.	x-coordinate, y-coordinate, rank/suit (e.g. 5H for five of hearts), location (0-15)
Game Information	Metadata concerning the game: difficulty of the game, seed to generate the deal, the starting time, and the end time of the game is logged	Easy/Normal/Hard, seed number, UNIX Timestamp

The general program flow can be seen in Fig. 3. First, the program waits for user events. Secondly, when these events arise, the program immediately captures the screen and corresponding user input. This combination is crucial to determine the action and outcome of the user. Thirdly, image processing is utilized to capture any visual cues of the outcome (e.g. the program cards that have been moved). Finally, the user input is combined with the visual cues of the game to evaluate the event. It is crucial that when such a game event happens, the program stops any calculations it is doing at that point (e.g. processing previous moves). To solve this, three threads are started at the beginning of the program. The first thread is the main thread, it processes inputs and screenshots, stores the digital biomarkers and handles other critical information such as coordinates of important locations. The second thread is purely dedicated to listening and capturing user input. The third thread is triggered by the user input and captures the next stable screenshot. This last thread has the highest priority of all threads, as capturing the screen as soon as possible after the user has made a move is crucial to determine the outcome.

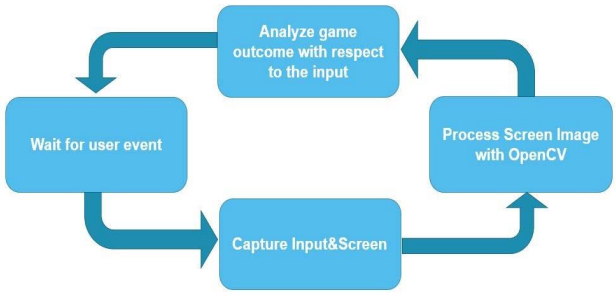

Fig. 3. General program flow

IV. Implementation

This section contains information concerning the implementation of the program. First, the setup is illustrated, explaining all necessary steps to play the game. Secondly, the card region extraction algorithm is clarified. Thirdly, how single cards are extracted is described. Fourthly, the card classification algorithm is explained. Finally, performance metrics of the program are given.

A. Setup

Once the program is activated, it starts the Microsoft Solitaire Collection. From there on, the program continuously monitors the state it is in: playing, choosing a game, selecting a difficulty, starting a game, ending a game, etc. To determine the state, the program follows the state diagram as illustrated in Fig. 4. Depending on the specific state of the game, interactions of the user will be interpreted in a different way. For example, in the PLAYING state, double-clicking will trigger an event to detect changes in the playing board state. While in the MAINMENU state double-clicking is ignored. To prevent essential board information from being obfuscated by pop-ups or animations, and to ensure move stability, hints need to be turned off in the settings, as well as single tap to move, alerts, tutorial, background animations, and end animations.

As some players may have the game open prior to launching the program, the game does an initial check on whether the starting state is MAINMENU or PLAYING. The central state of the program is the PLAYING state. Unless the player accesses the menu, requests hints, or starts a new game, the program will just follow the natural game progress. The program processes these changes in state by detecting button clicks. For these buttons, the dynamic position of the button is extracted through contour detection.

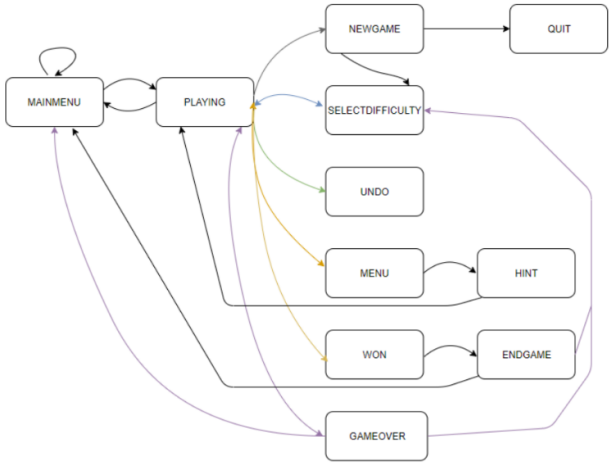

Fig. 4. State diagram of the program

B. Card Region Extraction

At the beginning of a game, each card is clearly visible. But as the game progresses, some card stacks tend to grow larger and cards tend to overlap (Fig. 5). Due to this overlap, each card needs to be extracted (i.e. the visible portion of each card needs to be separated) and classified (i.e. the rank and suit need to be determined) at the start of the game. This way, a model of the entire playing field is mapped. During the rest of the game, only the top cards are extracted and classified to monitor the progress of the game.

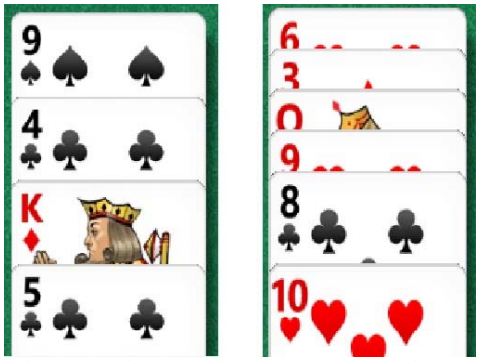

Fig. 5. At the start of the game, the rank and suit of each card is clearly visible (left). After the game progresses, the rank and suit can be hidden by overlapping cards (right)

To extract all cards, the screenshot of the board will be split into different pieces according to the card regions. For FreeCell, as aforementioned, there are eight regions on the *build stacks*, four on the *storage stacks* and four on the *suit stacks*. The coordinates of these 16 card regions will be used to determine the actual region the user interacts with. To find these regions, the screenshot of the board is first converted to grayscale, and next, by *thresholding* converted to a binary image (Fig. 6). Thresholding is an image segmentation technique that creates a binary image of a grayscale image based on a manually selected threshold. Then, all the contours of this image are found using the contour detection algorithms of OpenCV. Contours are the curves of continuous points that have the same color or intensity [67]. After filtering, only the contours that are larger than the size of a card (i.e. the card stacks) remain. We can draw a rectangle around each of these contours. The coordinates of these rectangles are then stored into a vector. This way, we can compare them with the coordinates of each click, thus being able to detect the cards the users interact with.

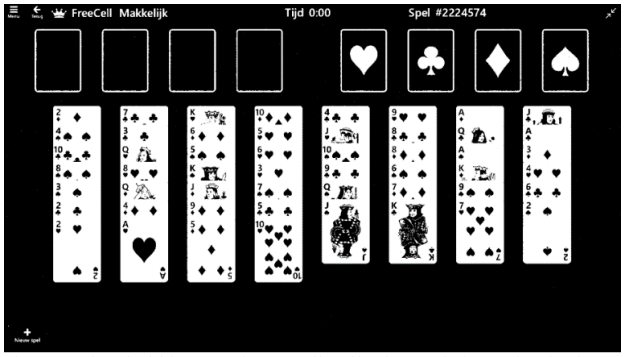

Fig. 6. A threshold image of FreeCell. All relevant information is shown in white while all noise is eliminated.

C. Unique Card Extraction

With the card regions defined, each card can be extracted and classified. The width of the cards is defined based on the width of the card region contours. As the ratio between the card width and height is resolution independent, the height can be inferred from the width. Then, the screenshot is divided into sixteen card regions as seen in Fig. 7.

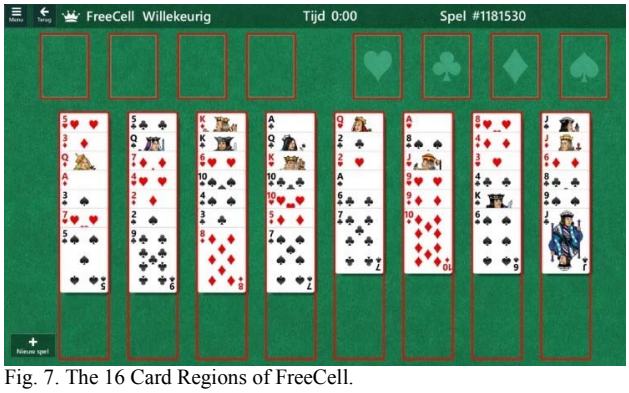

Fig. 7. The 16 Card Regions of FreeCell.

Since cards are stacked, not all are completely visible on the screen. At the start of the game, cards that are partially visible have an aspect ratio of 0.4 (width over length). This way, if we extract card images of 0.4 times the card height, we can extract the card sections with the rank and the suit clearly visible (Fig. 8). These sections are split in rank and suit using contour detection, they are stored in separate vectors and are ready for classification.

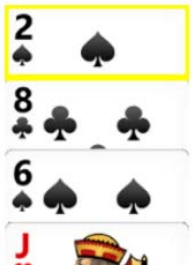

Fig. 8. An extracted card region with the first card extracted through the aspect ratio

D. Card Classification

To classify the rank and suit, three algorithms are applied, as shown in Fig. 9. First, the contours of the rank are extracted as individual images. Secondly, the images are rescaled to a standard size (40x50 pixels for ranks, 50x50 for suits). Finally, the images are converted into a binary black and white image. These binary images are classified using a k-nearest neighbors classifier trained on different sets of rank and suit images. With every rank and suit classified, a digital representation of the board can be build and consecutive player actions can be interpreted.

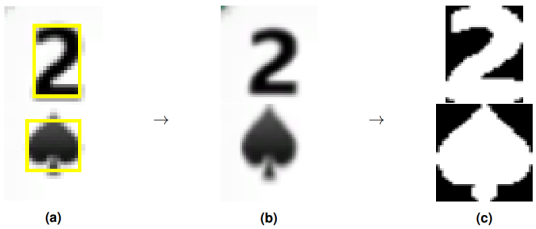

Fig. 9. (a) detecting the contours of the rank and suit (b) Extracted rank and suit. (c) The binary image of rank and suit ready for classification

E. Performance

To test the resource efficiency of our toolkit, a performance test was done on a computer (8GB RAM, i7 Intel Core 2.7 GHz). The extraction and classification performance of the whole board was tested over 10 different boards. Classification of the whole board, necessary at the start of the games, took on average 51ms. The classification of all top cards, necessary for interpreting moves, took on average 32.1ms. Finally, the extraction and classification of the game seed number on the top right took on average 260.3 ms. This performance allows for near real-time evaluation of gameplay, with response times that are lower than the theoretical limits of what can be perceived by players[68]. This was confirmed during the exploratory study (see chapter V). Participants of the exploratory tests did not notice any interference of the program while playing the game.

V. Exploratory Data Analysis

To explore the potential of the toolkit for capturing digital biomarkers, a first exploratory study was conducted. Digital biomarkers were captured from three different age groups. According to literature on cognitive performance, age-related cognitive decline is a natural process [69]–[71]. Primarily working memory, motor control, episodic memory, spatial ability, reasoning, and processing speed deteriorate as people grow older [71], [72]. In other words, people need more time to complete tasks and find it harder to keep important information in mind. Hence, the impact of age on player actions through digital biomarkers was explored as it may be indicative of cognitive performance. We aimed to explore whether digital biomarkers could discriminate among age groups, and possibly show a decline for the older groups. The biomarkers were categorized into three groups: Time-related, Error-related, and Outcome-related Digital Biomarkers.

A. Method

Digital biomarkers were captured from users across three different age groups (18-25,45-55,65+). The first age group, from now on referred to as youth, contained 21 participants. The middle group, referred to as middle-aged, contained 12 participants. The oldest group, referred to as elderly, contained 11 participants. Each of these participants lived independently, had no known cognitive impairments or prior cognitive complaints. In addition, all participants were new to FreeCell.

As they had no previous FreeCell experience, each participant was first briefed about the rules and mechanics of FreeCell via a fixed presentation. After this presentation, each participant got to play a practice game (seed number #25001) [73]. During this practice game, questions were allowed concerning the game rules. After this practice game, each participant played the same identical games (seeds #34898, #2365418 and #8840193). The choice for identical seeds eliminated differences in game performance due to the chance of having a more 'generous' deal. During these three games, questions were not allowed and players continued playing until they either finished the game, the game ended because of a lack of possible moves or until the user deemed that he/she was stuck and requested to end that game.

B. Results

This data was visually explored to give insight into age-related playing differences. The goal is to show the possibilities of the toolkit. The information is divided into three distinct categories: time-related digital biomarkers, outcome-related digital biomarkers, and error-related digital biomarkers.

1) Time-related Digital Biomarkers

We were most interested in the digital biomarkers related to time spent thinking before making a move as this can possibly correlate to important cognitive functions for daily activities such as attention, executive function, and planning. The 'young' age group has an average think-duration of 6871.84ms (sd: 2467.59 ms). The middle-aged group has an average think-duration of 10383.40ms (sd: 5816.19 ms), while the 'elderly' have an average think-duration of 13423.65ms (sd: 7089.50 ms) (Fig. 10).

We explored the difference in time spent thinking before an erroneous or successful move, as seen in Fig. 11 and Fig. 12. For successful moves, players in the youth category thought on average 6805.76 ms (sd: 2402.36 ms), players in the mid age category 10755.85 ms (sd: 5840.93 ms); and the oldest group 13241.61 ms (6750.04 ms). For erroneous moves, players in the youth category thought on average 7337.47 ms (sd: 4881.20 ms), players in the mid age category 8487.77 ms (sd: 5462.14 ms); and the oldest group 15448.09 ms (13395.10 ms). For all age groups, except for the mid age group, time spent thinking before making a successful move was shorter than for an erroneous move. The average think-duration of each move in time is shown in Fig. 16. The x-axis indicates the move number of the game, meaning, the first value on the x-axis corresponds to the first move of the game. The y-axis corresponds to the average think-duration of that specific move. Concerning Move Time, as shown in Fig. 13, people in the young category took on average 1578.82 ms (sd: 809.52) to move a card. For the middle-aged category this was 1661.28 ms (sd:791.49 ms). The oldest category took 2103.04 ms (sd: 1298.41 ms) on average to make a move.

2) Error-related Digital Biomarkers

Errors made during gameplay may be indicative of planning, executive functioning, and attention as players are required to think ahead, processing the next couple of moves. The average total amount of errors made by each age group van be seen in Fig. 14. On average, the youth group made 12.2 mistakes, the middle-aged group 12.4 mistakes, and the elderly made 7.1 mistakes. Concerning Rank Errors, the youngest group made 5.3 errors on average, the middle-aged group 5.0 errors, and the oldest group 3.7 errors. For Suit Errors, 3.5, 3.6, and 1.6 errors were made on average for the youngest, middle, and oldest category. In regard to Unmovable Card Errors, the youngest, middle aged, and oldest category made 1.2, 1.3, and 0.5 errors respectively. Regarding the Too Many Cards Moved errors, the youngest age group made 2.3 errors on average, the middle-aged group 4.3 errors and the eldest group 2.5 errors. For requesting hints an average of 0.04 was found for the youngest group, 0.25 for the middle-aged group, and 1.00 for the oldest group. For correcting unwanted moves, on average, 0.86 was found for the youngest group, 0.31 for the middle-aged group, and 0.39 for the oldest group.

3) Outcome-related Digital Biomarkers

Fig. 15 shows the percentages of the game won and lost. The results display that the win rate decreases with the increase of age group. The percent of games won by youth, mid age and elderly is, 91.3%, 77.1% and 60.9% respectively.

VI. Discussion

Today, many cognitive impairments go undiagnosed, and those patients who have been diagnosed have sparse follow-up moments with neuropsychologists, due to restricted time and funding. This problem may be mitigated by adding digital biomarkers to the toolbox of neuropsychological assessment. Previous research has shown that daily interactions with technology can provide a trail of information on cognitive performance. This may be an efficient method of gathering digital biomarkers whilst reducing the effort from the user and

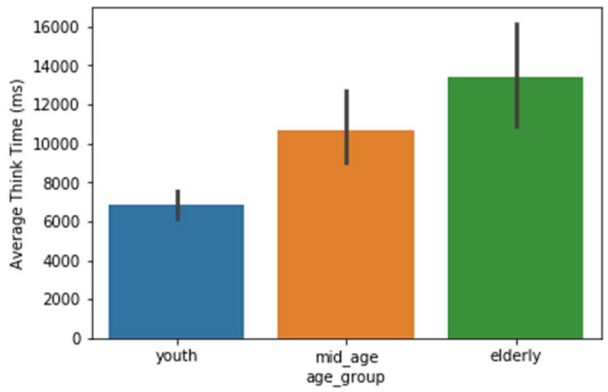

Fig. 10. Average Think Time. The vertical line resembles the 95% confidence interval

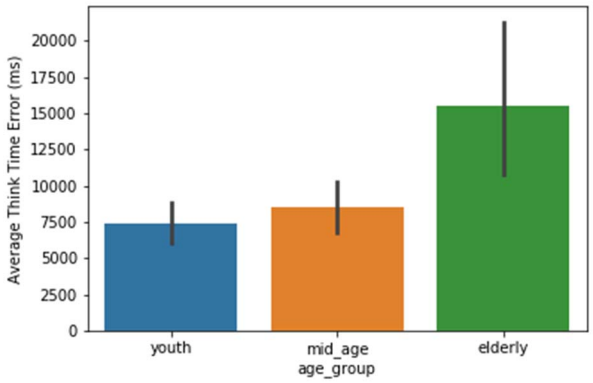

Fig. 11. Average Think Time for an error

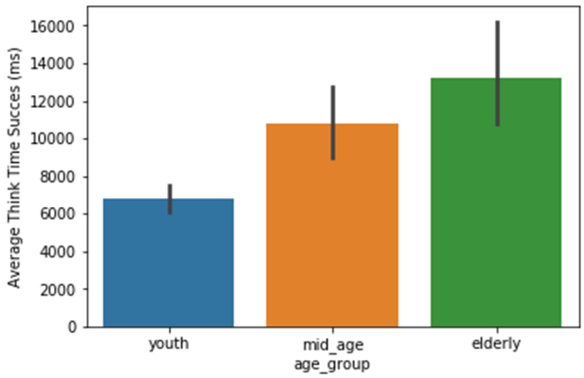

Fig. 12. Average Think Time for a successful move

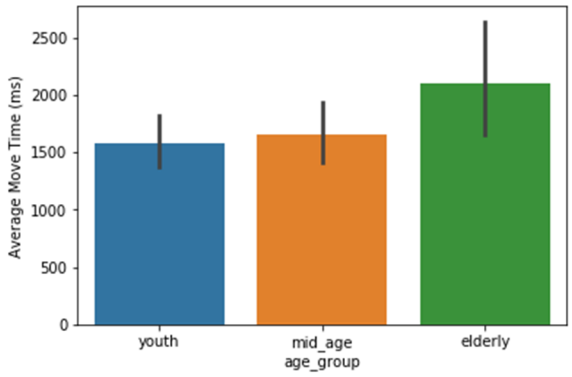

Fig. 13. Average Move Time

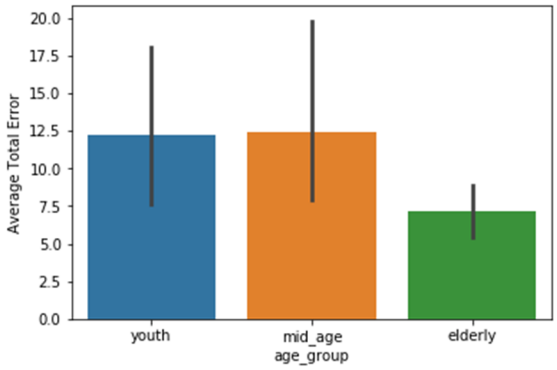

Fig. 14. Average Total Errors

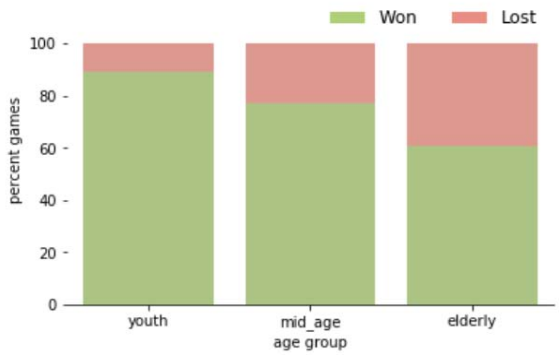

Fig. 15. Total Percentage of games won

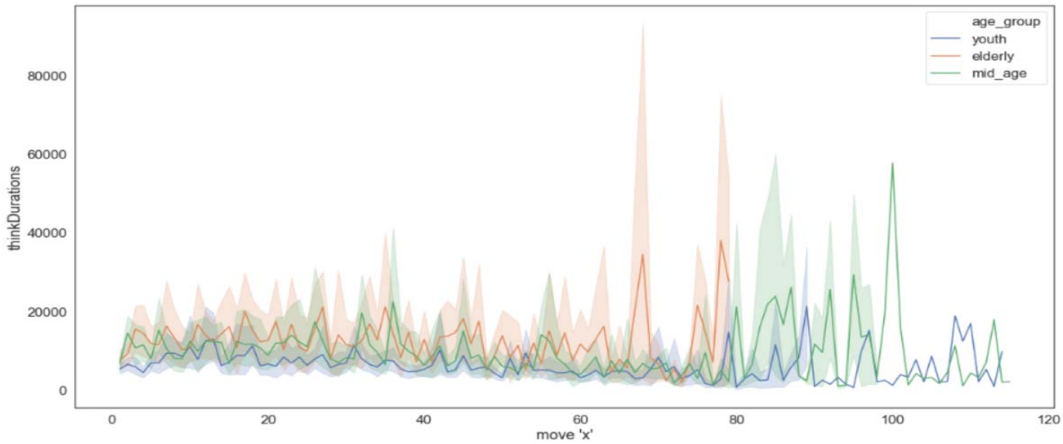

Fig. 16. Average Think Time on the x^{th} move

the healthcare system. Such digital biomarkers can help fill in the gap between consultations and can potentially help in screening, diagnosis, and prognosis of cognitive health.

In this paper, we explored a novel way of capturing digital biomarkers via gameplay of card games, by utilizing image processing. The contribution of our work lies in the presentation of the generic toolkit using image processing, and a first exploration of player actions that can be indicative of cognitive performance.

We built a generic image processing toolkit for card games to collect digital biomarkers on cognitive health. We chose the Microsoft Solitaire Collection, coded general image processing algorithms to detect cards, and implemented game rules for two games: Klondike Solitaire and FreeCell. As these card games are popular amongst young and old, and weaved in the daily lives of people, it may be a good fit to capture digital biomarkers. Modularity of the program was kept in mind while programming, such that this toolkit can easily be extended to other card games such as TriPeaks and Pyramid. The card detection algorithm is generalizable for all card games in the Microsoft Solitaire set, only the game rules need to be implemented. Performance tests showed that this toolkit is able to capture digital biomarkers at real-time with minimum stress on the computer. All participants of the exploratory tests did not notice any interference of the program while playing the game. No visual or auditory queues came up. The program did not stress the performance of the computer, gameplay remained as smooth as if the toolkit's software program was not there. However, technical improvements can still be made for the toolkit. Up to date, threshold values are manually selected. This can be set automatically using techniques such as Otsu Thresholding, making the program more resilient to changes [74]. Furthermore, currently, all animations need to be turned off to ensure image stability during processing. Hence, updates could be made to make the toolkit more durable and less susceptible to animations. Moreover, more advanced machine learning techniques, such as deep learning, can be explored to further improve the robustness of classification [75]–[77]. The toolkit can also be adapted to other 2D-games with minor adjustments. New game rules need to be implemented and the machine learning models need to be retrained to detect new targets. Furthermore, new digital biomarkers need to be defined as not all digital biomarkers from this study are generalizable for all games.

As a first exploratory study, data from 44 participants from three different age groups was captured. In this paper, we limited us to visualizing the data and descriptive statistics. Results from this exploratory study suggest that data gathered via the generic toolkit can discriminate among different age groups of cognitively healthy participants, and possibly provide information on cognitive performance. At a group level, all time-related digital biomarkers show a steady decline the older the age group. This can be expected as cognitive functions critical to cognitive aging such as processing speed and working memory tend to decline[78]. As expected, the older the age group, the less games were won on average. However, for Error-related digital biomarkers, the reverse was true. Older adults made less errors than their younger and middle aged counterparts. This could indicate that older adults need more time to think of a move, but make their moves with more caution.

However, as this is a first, exploratory investigation, this study also has its shortcomings and any interpretations need to be done in a conditional manner. First, the groups were small and unbalanced, making results not generalizable to a wider population. Secondly, we compared cognitive healthy age groups as opposed to groups with cognitive impairments. Thirdly, differences were found at the group level only, no investigation was carried out at the individual level. If daily interactions are going to be predictors of cognitive performance, results should be obtained at the individual level. To this end, data should be captured over a longer period of time. Results should be compared inter-group and intra-individually. In this manner, a more accurate analysis of digital biomarkers as bearers of cognitive information can be performed (i.e. improving sensitivity and specificity). Hence, in the future, data should be captured over a longer period of time from larger populations, and populations with cognitive impairments. Ultimately, with the help of *machine learning models*, cognitively healthy participants could be discerned from their impaired counterparts, on the basis of multiple combined digital biomarkers. Finally, more complex composite digital biomarkers should be explored.

VII. Conclusion

Early diagnosis and frequent follow-up of cognitive health problems is crucial to managing disease progression, allowing for timely treatment. Digital biomarkers obtained via gameplay have potential to aid in early diagnosing of cognitive health issues. To this end, we developed a generic toolkit for card games, using image processing, to capture digital biomarkers indicative of cognitive performance. First, we applied a methodical approach to define 10 digital biomarkers indicative of cognitive performance. Next, we implemented the toolkit, on the top of the Microsoft 10 Solitaire Collection, as a multithreaded C++ desktop application, utilizing the Open Source Computer Vision Library to unobtrusively monitor games. Performance tests showed that this toolkit is able to capture digital biomarkers at real-time with minimum stress on the CPU. Finally, we conducted an exploratory user study to verify whether we can discriminate amongst different age groups, characterized by different cognitive performance due to normal cognitive ageing. The results of the exploratory study suggest, at a group level, that age groups differ. Time-based digital biomarkers and outcome related measures show a steady decline the older the age group. Although this is only a first exploratory study, the results suggest promise of the use of games, weaved in the daily life of players, for the capturing of digital biomarkers for cognitive health.

VIII. SOURCECODE

We would like to invite all researchers to build on, repurpose, and utilize this tool. All source code can be found on https://github.com/kgielis/ImageProcessingMicrosoftSolitaireCollection. All work based on this code should be referenced correctly. Fair use and modification is allowed, as described by The GNU General Public License v3.0.

IX. References

[1] "WHO | Dementia: a public health priority," *WHO*. [Online]. Available: http://www.who.int/mental_health/neurology/dementia/en/. [Accessed: 31-Jan-2019].

[2] "Depression and Other Common Mental Disorders: Global Health Estimates," World Health Organization, Geneva, 2017.

[3] "Comprehensive mental health action plan 2013–2020," World Health Organisation, sixty-sixth world health assembly, May 2013.

[4] D. Chisholm *et al.*, "Scaling-up treatment of depression and anxiety: a global return on investment analysis," *The Lancet Psychiatry*, vol. 3, no. 5, pp. 415–424, May 2016.

[5] R. C. Petersen *et al.*, "Current concepts in mild cognitive impairment," *Arch. Neurol.*, vol. 58, no. 12, pp. 1985–1992, Dec. 2001.

[6] Alzheimer's Association, "2016 Alzheimer's Disease Facts and Figures," Alzheimer's & Dementia 2016.

[7] G. Mushtaq, N. H. Greig, J. A. Khan, and M. A. Kamal, "Status of acetylcholinesterase and butyrylcholinesterase in Alzheimer's disease and type 2 diabetes mellitus," *CNS Neurol Disord Drug Targets*, vol. 13, no. 8, pp. 1432–1439, 2014.

[8] C. Laske *et al.*, "Innovative diagnostic tools for early detection of Alzheimer's disease," *Alzheimer's & Dementia*, vol. 11, no. 5, pp. 561–578, May 2015.

[9] J. W. Feightner and G. Worrall, "Early detection of depression by primary care physicians," *CMAJ*, vol. 142, no. 11, pp. 1215–1220, Jun. 1990.

[10] G. M. Savva and A. Arthur, "Who has undiagnosed dementia? A cross-sectional analysis of participants of the Aging, Demographics and Memory Study," *Age Ageing*, vol. 44, no. 4, pp. 642–647, Jul. 2015.

[11] "Biomarkers In Risk Assessment: Validity And Validation (EHC 222, 2001)." [Online]. Available: http://www.inchem.org/documents/ehc/ehc/ehc222.htm. [Accessed: 01-Feb-2019].

[12] "Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework," *Clinical Pharmacology & Therapeutics*, vol. 69, no. 3, pp. 89–95, 2001.

[13] N. Sharma and A. N. Singh, "Exploring Biomarkers for Alzheimer's Disease," *J Clin Diagn Res*, vol. 10, no. 7, pp. KE01–KE06, Jul. 2016.

[14] Z. S. Nasreddine *et al.*, "The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment: MOCA: A BRIEF SCREENING TOOL FOR MCI," *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, Apr. 2005.

[15] T. N. Tombaugh and N. J. McIntyre, "The Mini-Mental State Examination: A Comprehensive Review," *Journal of the American Geriatrics Society*, vol. 40, no. 9, pp. 922–935, Sep. 1992.

[16] R. M. Ahmed *et al.*, "Biomarkers in dementia: clinical utility and new directions," *J Neurol Neurosurg Psychiatry*, vol. 85, no. 12, pp. 1426–1434, Dec. 2014.

[17] K. Blennow, H. Hampel, M. Weiner, and H. Zetterberg, "Cerebrospinal fluid and plasma biomarkers in Alzheimer disease," *Nature Reviews Neurology*, vol. 6, no. 3, pp. 131–144, Mar. 2010.

[18] J. D. Doecke *et al.*, "Blood-Based Protein Biomarkers for Diagnosis of Alzheimer Disease," *Arch Neurol*, vol. 69, no. 10, pp. 1318–1325, Oct. 2012.

[19] C. Reitz and R. Mayeux, "Alzheimer disease: Epidemiology, diagnostic criteria, risk factors and biomarkers," *Biochemical Pharmacology*, vol. 88, no. 4, pp. 640–651, Apr. 2014.

[20] T. Smith, N. Gildeh, and C. Holmes, "The Montreal Cognitive Assessment: Validity and Utility in a Memory Clinic Setting," *Can J Psychiatry*, vol. 52, no. 5, pp. 329–332, May 2007.

[21] T. Tapiola *et al.*, "Cerebrospinal Fluid β-Amyloid 42 and Tau Proteins as Biomarkers of Alzheimer-Type Pathologic Changes in the Brain," *Arch Neurol*, vol. 66, no. 3, pp. 382–389, Mar. 2009.

[22] T. N. Tombaugh and N. J. McIntyre, "The Mini-Mental State Examination: A Comprehensive Review," *Journal of the American Geriatrics Society*, vol. 40, no. 9, pp. 922–935, 1992.

[23] "Digital Biomarkers | Karger Journal."

[24] R. L. Mandryk and M. V. Birk, "The Potential of Game-Based Digital Biomarkers for Modeling Mental Health," *JMIR Ment Health*, vol. 6, no. 4, p. e13485, Apr. 2019.

[25] S. A. Cooley *et al.*, "Longitudinal Change in Performance on the Montreal Cognitive Assessment in Older Adults," *The Clinical Neuropsychologist*, vol. 29, no. 6, pp. 824–835, Aug. 2015.

[26] B. De Schutter, "Never Too Old to Play: The Appeal of Digital Games to an Older Audience," *Games and Culture*, vol. 6, no. 2, pp. 155–170, Mar. 2011.

[27] H. Jimison, M. Pavel, J. McKanna, and J. Pavel, "Unobtrusive monitoring of computer interactions to detect cognitive status in elders," *IEEE Trans Inf Technol Biomed*, vol. 8, no. 3, pp. 248–252, Sep. 2004.

[28] J. L. Palmer, M. A. Coats, C. M. Roe, S. M. Hanko, C. Xiong, and J. C. Morris, "Unified Parkinson's Disease Rating Scale-Motor Exam: inter-rater reliability of advanced practice nurse and neurologist assessments: Neurological assessment inter-rater reliability," *Journal of Advanced Nursing*, vol. 66, no. 6, pp. 1382–1387, Apr. 2010.

[29] S. K. Inouye, C. H. van Dyck, C. A. Alessi, S. Balkin, A. P. Siegal, and R. I. Horwitz, "Clarifying confusion: the confusion assessment method. A new method for detection of delirium," *Ann. Intern. Med.*, vol. 113, no. 12, pp. 941–948, Dec. 1990.

[30] "The measure of you." [Online]. Available: https://www.roche.com/about/priorities/personalised_healthcare/digital-biomarkers.htm. [Accessed: 14-Feb-2019].

[31] E. R. Dorsey, S. Papapetropoulos, M. Xiong, and K. Kieburtz, "The First Frontier: Digital Biomarkers for Neurodegenerative Disorders," *Digital Biomarkers*, Jul. 2017.

[32] J. Torous, J.-P. Onnela, and M. Keshavan, "New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices," *Transl Psychiatry*, vol. 7, no. 3, p. e1053, 07 2017.

[33] A. Shehab and A. Abdulle, "Cognitive and autonomic dysfunction measures in normal controls, white coat and borderline hypertension," *BMC Cardiovascular Disorders*, vol. 11, no. 1, p. 3, Jan. 2011.

[34] M. M. Redfield *et al.*, "Isosorbide Mononitrate in Heart Failure with Preserved Ejection Fraction," *New England Journal of Medicine*, vol. 373, no. 24, pp. 2314–2324, Dec. 2015.

[35] S. Saeb *et al.*, "Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study," *Journal of Medical Internet Research*, vol. 17, no. 7, p. e175, Jul. 2015.

[36] M. Faurholt-Jepsen, M. Vinberg, M. Frost, E. M. Christensen, J. E. Bardram, and L. V. Kessing, "Smartphone data as an electronic biomarker of illness activity in bipolar disorder," *Bipolar Disorders*, vol. 17, no. 7, pp. 715–728, Nov. 2015.

[37] T. Beiwinkel *et al.*, "Using Smartphones to Monitor Bipolar Disorder Symptoms: A Pilot Study," *JMIR Mental Health*, vol. 3, no. 1, p. e2, Jan. 2016.

[38] I. Barnett, J. Torous, P. Staples, L. Sandoval, M. Keshavan, and J.-P. Onnela, "151. Relapse Prediction in Schizophrenia through Digital Phenotyping," *Biological Psychiatry*, vol. 83, no. 9, pp. S61–S62, May 2018.

[39] S. Hagler, D. Austin, T. L. Hayes, J. Kaye, and M. Pavel, "Unobtrusive and Ubiquitous In-Home Monitoring: A Methodology for Continuous Assessment of Gait Velocity in Elders," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 813–820, Apr. 2010.

[40] B. Suits, T. Hurka, and F. Newfeld, *The grasshopper: games, life, and utopia*, 3. ed. Peterborought, Ont: Broadview Press, 2014.

[41] J. McGonigal, *Reality is broken: why games make us better and how they can change the world*. New York: Penguin Press, 2011.

[42] M. Csikszentmihalyi, *Finding flow: The psychology of engagement with everyday life*. New York, NY, US: Basic Books, 1997.

[43] J. DeRight and R. S. Jorgensen, "I Just Want My Research Credit: Frequency of Suboptimal Effort in a Non-Clinical Healthy Undergraduate Sample," *The Clinical Neuropsychologist*, vol. 29, no. 1, pp. 101–117, Jan. 2015.

[44] S. Valladares-Rodriguez, M. J. Fernández-Iglesias, L. Anido-Rifón, D. Facal, and R. Pérez-Rodríguez, "Episodix: a serious game to detect cognitive impairment in senior adults. A psychometric study," *PeerJ*, vol. 6, p. e5478, Sep. 2018.

[45] P. Słowiński *et al.*, "Unravelling socio-motor biomarkers in schizophrenia," *npj Schizophrenia*, vol. 3, no. 1, Dec. 2017.

[46] H.-T. Jung *et al.*, "Estimating Mini Mental State Examination Scores using Game-Specific Performance Values: A Preliminary Study," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, 2018, pp. 1518–1521.

[47] K. Leduc-McNiven, B. White, H. Zheng, R. D McLeod, and M. R Friesen, "Serious games to assess mild cognitive impairment: 'The game is the assessment,'" *Research and Review Insights*, vol. 2, no. 1, 2018.
[48] C. Boletsis and S. McCallum, "Smartkuber: A Serious Game for Cognitive Health Screening of Elderly Players," *Games for Health Journal*, vol. 5, no. 4, pp. 241–251, Aug. 2016.
[49] P. Felicia, Ed., *Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches*. IGI Global, 2011.
[50] M. Colder Carras *et al.*, "Commercial Video Games As Therapy: A New Research Agenda to Unlock the Potential of a Global Pastime," *Frontiers in Psychiatry*, vol. 8, Jan. 2018.
[51] J. Lumsden, E. A. Edwards, N. S. Lawrence, D. Coyle, and M. R. Munafò, "Gamification of Cognitive Assessment and Cognitive Training: A Systematic Review of Applications and Efficacy," *JMIR Serious Games*, vol. 4, no. 2, p. e11, Jul. 2016.
[52] P. L. Ackerman, R. Kanfer, and C. Calderwood, "Use it or lose it? Wii brain exercise practice and reading for domain knowledge.," *Psychology and Aging*, vol. 25, no. 4, pp. 753–766, 2010.
[53] A. C. S. Torres, "Cognitive effects of video games on old people," *International Journal on Disability and Human Development*, vol. 10, no. 1, Jan. 2011.
[54] K. Salen and E. Zimmerman, "Game design and meaningful play," *Handbook of Computer Game Studies*, pp. 59–79, 2005.
[55] W. R. Boot, D. Souders, N. Charness, K. Blocker, N. Roque, and T. Vitale, "The Gamification of Cognitive Training: Older Adults' Perceptions of and Attitudes Toward Digital Game-Based Interventions," in *Human Aspects of IT for the Aged Population. Design for Aging*, vol. 9754, J. Zhou and G. Salvendy, Eds. Cham: Springer International Publishing, 2016, pp. 290–300.
[56] O. Thompson, S. Barrett, C. Patterson, and D. Craig, "Examining the Neurocognitive Validity of Commercially Available, Smartphone-Based Puzzle Games," *Psychology*, vol. 03, no. 07, pp. 525–526, 2012.
[57] M. Pavel, H. Jimison, S. Hagler, and J. McKanna, "Using Behavior Measurement to Estimate Cognitive Function Based on Computational Models," in *Cognitive Informatics in Health and Biomedicine: Understanding and Modeling Health Behaviors*, V. L. Patel, J. F. Arocha, and J. S. Ancker, Eds. Cham: Springer International Publishing, 2017, pp. 137–163.
[58] J. W. Grabbe, "Sudoku and Working Memory Performance for Older Adults," *Activities, Adaptation & Aging*, vol. 35, no. 3, pp. 241–254, Jul. 2011.
[59] "OpenCV library." [Online]. Available: http://opencv.org/. [Accessed: 30-Aug-2017].
[60] A. Elyasaf, A. Hauptman, and M. Sipper, "Evolutionary Design of FreeCell Solvers," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 4, no. 4, pp. 270–281, Dec. 2012.
[61] P. L. Baker, "Using FreeCell to teach mathematics," *The Mathematics Teacher; Reston*, vol. 96, no. 6, pp. 406–410, Sep. 2003.
[62] H. Jimison, M. Pavel, J. McKanna, and J. Pavel, "Unobtrusive Monitoring of Computer Interactions to Detect Cognitive Status in Elders," *IEEE Transactions on Information Technology in Biomedicine*, vol. 8, no. 3, pp. 248–252, Sep. 2004.
[63] M. W. H. Chen, "The Hardness Analysis of FreeCell Games," p. 33.
[64] R. Jilani, A. Crampton, D. Kitchin, and M. Vallati, "Have a Little Patience: Let Planners Play Cards," p. 7.
[65] A. Elyasaf, A. Hauptman, and M. Sipper, "GA-FreeCell: evolving solvers for the game of FreeCell," in *Proceedings of the 13th annual conference on Genetic and evolutionary computation - GECCO '11*, Dublin, Ireland, 2011, p. 1931.
[66] A. Elyasaf, A. Hauptman, and M. Sipper, "Evolutionary Design of FreeCell Solvers," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 4, no. 4, pp. 270–281, Dec. 2012.
[67] "OpenCV: Contours : Getting Started." [Online]. Available: https://docs.opencv.org/3.4.3/d4/d73/tutorial_py_contours_begin.html. [Accessed: 17-Jan-2019].
[68] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *Nature*, vol. 381, no. 6582, p. 520, Jun. 1996.
[69] T. Hedden and J. D. E. Gabrieli, "Insights into the ageing mind: a view from cognitive neuroscience," *Nature Reviews Neuroscience*, vol. 5, no. 2, pp. 87–96, Feb. 2004.
[70] J. H. Morrison and M. G. Baxter, "The ageing cortical synapse: hallmarks and implications for cognitive decline," *Nature Reviews Neuroscience*, vol. 13, no. 4, pp. 240–250, Apr. 2012.
[71] N. A. Bishop, T. Lu, and B. A. Yankner, "Neural mechanisms of ageing and cognitive decline," *Nature*, vol. 464, pp. 529–535, Mar. 2010.
[72] R. D. Seidler *et al.*, "Motor Control and Aging: Links to Age-Related Brain Structural, Functional, and Biochemical Effects," *Neurosci Biobehav Rev*, vol. 34, no. 5, pp. 721–733, Apr. 2010.
[73] "Get Microsoft Solitaire Collection," *Microsoft Store*. [Online]. Available: https://www.microsoft.com/en-us/p/microsoft-solitaire-collection/9wzdncrfhwd2. [Accessed: 21-Feb-2019].
[74] X. Xu, S. Xu, L. Jin, and E. Song, "Characteristic analysis of Otsu threshold and its applications," *Pattern Recognition Letters*, vol. 32, no. 7, pp. 956–961, May 2011.
[75] S. Zhou, Q. Chen, and X. Wang, "Discriminative Deep Belief Networks for image classification," in *2010 IEEE International Conference on Image Processing*, Hong Kong, Hong Kong, 2010, pp. 1561–1564.
[76] Y. Tang, J. Wang, B. Gao, E. Dellandrea, R. Gaizauskas, and L. Chen, "Large Scale Semi-Supervised Object Detection Using Visual and Semantic Knowledge Transfer," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 2119–2128.
[77] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: A Simple Deep Learning Baseline for Image Classification?," *IEEE Trans. on Image Process.*, vol. 24, no. 12, pp. 5017–5032, Dec. 2015.
[78] C. N. Harada, M. C. Natelson Love, and K. L. Triebel, "Normal Cognitive Aging," *Clinics in Geriatric Medicine*, vol. 29, no. 4, pp. 737–752, Nov. 2013.