

A Cost-Effective Eye-Tracker for Early Detection of Mild Cognitive Impairment

Danilo Greco¹ Francesco Masulli^{1 2} Stefano Rovetta^{1 2} Alberto Cabri^{1 2} Davide Daffonchio¹

¹DIBRIS University of Genoa - Via Dodecaneso, 35 - 16146 Genoa (Italy)

²Vega Research Laboratories s.r.l. - Via Ippolito d'Aste, 7/5 - 16121 Genoa (Italy)
francesco.masulli@unige.it

Abstract—This paper presents a low-cost eye-tracker aimed at carrying out tests based on a Visual Paired Comparison protocol for the early detection of Mild Cognitive Impairment. The proposed eye-tracking system is based on machine learning algorithms, a standard webcam, and two personal computers that constitute, respectively, the "Measurement Sub-System" performing the test on the patients and the "Test Management Sub-System" used by medical staff for configuring the test protocol, recording the patient data, monitoring the test and storing the test results. The system also integrates an stress estimator based on the measurement of heart rate variability obtained with photoplethysmography.

Index Terms—Eye-tracker, Alzheimer's disease, Mild Cognitive Impairment, Early Detection, Raspberry, Webcam, Heart Rate Variability, Photoplethysmography, Python, Neural Networks

I. INTRODUCTION

Eye-trackers are devices designed to measure the direction of the gaze (where one is looking) and/or both eye movements related to the head. Nowadays, they are widely used in many fields, ranging from neuroscience, marketing, product design and human-machine interaction [1].

There are several methods for measuring eye movements, including search coils [2] and electrooculograms [3], but currently, the most popular eye-trackers use video images from which the eye position is extracted [4], [5].

Some cognitive impairments, such as Alzheimer's Disease (AD), are linked with eye movement-related disorders, while others indirectly relate to the gaze dynamics [6]. Several investigations have been conducted in the last few years aimed at applying eye-tracking technology to the detection of AD.

The main objective of medical research on AD is to diagnose the disease as soon as possible to maximize the effectiveness of therapies [7]. It is even possible to identify this disorder before symptoms appear by detecting Mild Cognitive Impairment (MCI). Patients with symptoms of MCI have a very high risk of developing AD, with a conversion rate of 6% to 25% per year [8]. Although many patients with MCI tend to become stable after some time, more than half of them regress into dementia within five years, which outlines MCI as an excellent alarm bell that anticipates AD development.

Even if MCI does not markedly impair the daily activities of a person, some symptoms consist in difficulties in carrying out complex operations that previously took place without problems (such as preparing a meal) [9]. It is therefore extremely important to identify symptoms of MCI from its early stages. Diagnostic biomarkers able to identify the pathology associated with this disorder include the detection of altered levels of tau and amyloid in cerebrospinal fluid, the use of structural Magnetic Resonance Imaging (MRI) to identify disease-specific patterns of regional atrophy and MRI T1 ρ to detect disease-related macro-molecular protein aggregation, and the direct imaging of amyloid deposits using positron emission tomography and single photon emission computerized tomography [10].

As it is not possible to foresee a wide diffusion of all those diagnostic tests to the entire adult population, given their high costs and degree of invasiveness, eye-tracking-based tests are very promising in order to obtain reliable, low-cost and non-invasive biomarkers for the early detection of MCI [11]. In fact, there are at least two types of tests based on eye-tracking that allow us to obtain biomarkers for MCI:

- *Saccadic Eye Movements (SEM) analysis*

Saccades are rapid eye movements of the eye followed by fixations, that is, the time in which the look is still to send information to the brain and proceed to the next saccade. These movements are necessary and customarily used to perform actions such as reading. Patients with AD presents irregularities of the saccades and show difficulties reading the texts. These disorders are already present in a very early stage of AD and in MCI before AD regression [12], [13].

- *Visual Paired Comparison (VPC)*

This type of test [7] relies on memory impairment caused by MCI. The VPC tests may consist into two distinct phases: the familiarization phase and the test phase. Both test steps display pairs of images for a assigned time on a screen, using the eye-tracker to monitor the observer's fixations. As an example, the pictures are presented in pairs with identical stimulus during the familiarization phase, while during the

test phase, the pairs of images to be displayed are created by placing on one side of the screen a known image already seen in the familiarization phase and on the other a new image. Healthy people usually observe the new one 70% of the stimulus duration. Patients with MCI, on the other hand, do not give preference to either of the two stimuli, reinforcing the hypothesis that they do not remember the images of the familiarization phase. These data highlight the short-term memory damage caused by MCI [14]. VPC tests are versatile and simple and they have been used with rodents, primates, children, and adults [15]. [16] showed that the VPC tests are able to diagnose brain damage caused by MCI not evidenced by MRI.

We note that a test for the early diagnosis of Mild Cognitive Impairment based on Visual Paired Comparison requires the measurement of the duration of the fixations on the images presented on the left and on the right side of a screen. As fixations in eye movements are events lasting at least 100 ms [17], it is therefore possible to use for eye-tracking a standard webcam with a frame rate of 30 fps.

Starting from this consideration, we designed an eye-tracking system based on two computers and a standard webcam that implements a Visual Paired Comparison for the early diagnosis of the Mild Cognitive Impairment. The two computers constitute, respectively, the "Measurement Sub-System" and the "Test Management Sub-System". The "Measurement Sub-System" consists of a desktop computer with a display of at least 24" with at least a I3 CPU and a webcam (or, better, an all-in-one computer with the same characteristics), or, also, a single-board computer such as a Raspberry (see Fig. 1) attached to a webcam and a display of at least 24". The observer to whom the test is administered is positioned in front of the display. The "Test Management Sub-System" is a desktop



Fig. 1. Raspberry Pi 4 and Raspberry Pi Camera Module 2.

or a laptop with at least a I3 CPU and is used by medical staff for configuring the test protocol, recording the patient data, monitoring the test and storing the test results.

The rest of this paper is organized as follows: Sec. II presents the technical aspects of the system we designed and the evaluation of its performance; Sec. III contains an example of protocol of a test for the early diagnosis of Mild Cognitive Impairment based on Visual Paired Comparison; Discussion and Conclusions are depicted in Sec. IV.

II. SYSTEM IMPLEMENTATION

A. System Design

Our approach to the eye-tracking design is based on the usage of video images from a webcam from which the gaze point on the screen is estimated [4], [5].

The necessary steps for gaze point measurement are the detection the position of the observer's face inside the images acquired by the webcam and then the position of the eyes into them and the estimation of the observed point on the screen.

The software was developed using the Python language [18] under the Linux operating system. It is based on two classifiers exploiting the Viola-Jones object detection framework [19], [20], the former used to detect the position of the observer's face and the latter used to detect the position of the eyes and on a dense neural network used for calibration [21]. The Viola-Jones object detection framework is a robust and real-time algorithm primarily motivated by the problem of face detection that is composed by four stages:

- 1) Haar Feature Selection, using Haar basis functions [22];
- 2) Creating an Integral Image [23], [24];
- 3) Adaboost Training [25];
- 4) Cascading Classifiers.

The Haar feature-based cascade classifiers [19], [20] are provided by a function of the OpenCV library that is a large open-source library for computer vision that we use also for images and webcam frames manipulation and for the implementation of the graphical user interfaces [26],

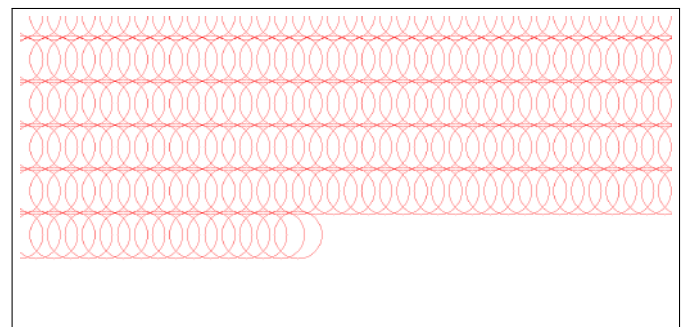


Fig. 2. Circle following a predefined pattern, used for calibration.

together with the PyQT library [27], while the neural network is implemented using Tensorflow [28].

B. Calibration procedure and assesment

Before the actual test it is necessary to carry out for each observer a calibration procedure where the geometric characteristics of the observer's eyes are (implicitly) estimated as the basis for a fully-customized and accurate gaze point prediction.

During calibration, a blank screen is presented to the observer with a circle that moves following a predefined pattern (Fig. 2). The observer is requested to keep his face still, with eyes opened and visible, the face frontal and centered with respect to the camera and to pursuit the circle with the eyes; moreover, only a single face must appear in the frame. If one of these conditions are not met, the program gives an alarm and goes into stand-by, waiting for the observer to position himself correctly again.

The eye images and the effective positions of the stimuli on the screen are collected in a data set that is used for training a dense neural network [29] aimed at creating a (in general non-linear) map that will allow to estimate the positions of the gaze points on the screen during the test.

The architecture of the dense neural network comprises an input layer of 1024 neurons acting on the sub-image 32x32 containing the eye, two hidden layers of 500 and 100 ReLU units and an output layer composed by one sigmoidal neuron giving the estimate horizontal position of the observed point on the screen. As the training is aimed at interpolation, we bring the optimization procedure into over-fitting (using a large number of epochs) and then obtain a precise localization of the observed point on the screen.

To quantitatively measure the accuracy of the neural network, a test procedure has been developed in which the observer is required to click on points that appear one-by-one in random positions on the right or left side of the screen. Based on the answers given by the observer, the program gives an evaluation of the accuracy on the right / left side discrimination (see Fig. 3).

The test of accuracy on 100 points (50 on the left and 50 on the right of the screen) shows an average absolute

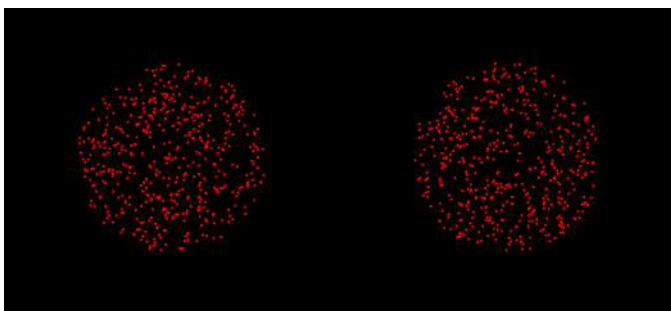


Fig. 3. Click-points pattern used to measure the system's accuracy.

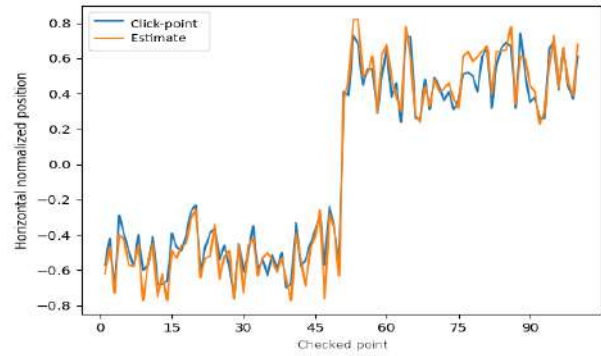


Fig. 4. Accuracy evaluation: comparison of the horizontal position of the click-point vs the neural network estimate.

error of 3.00% (1.6 cm on a 24-inch screen), with an ability to distinguish the right from the left of 100% (see Fig. 4).

C. Heart Rate Variability Monitoring

It's worth noting that it is possible to exploit the proposed architecture to add additional functionalities useful for monitoring the mental state of the patient during the MCI test, such as facial expression detection [30], and stress monitoring through the measurement of heart rate [31] or heart beats per minute (BPM).

In our system, a region of interest (ROI) Ω_{ROI} located in the forehead is selected to estimate the heart rate [31]. This is done by measuring in the sRGB color space the variation of the average optical intensity in the region in the e green channel that contains the most pulse

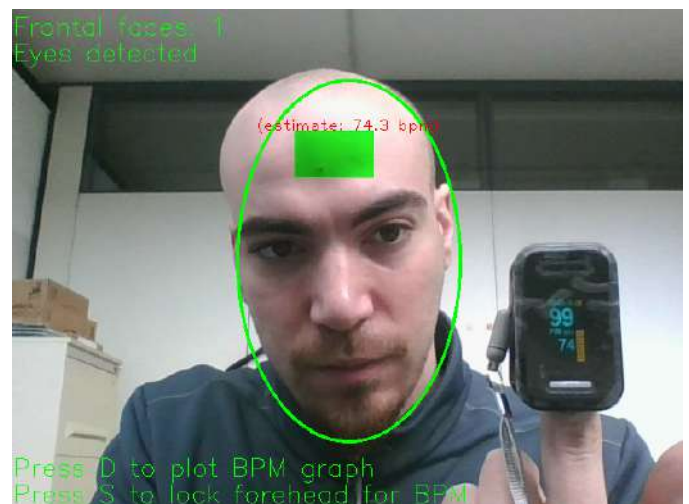


Fig. 5. Calibration phase: the ellipse turns green if there is a face inside and the eyes are detected. BPM is also estimated. This figure compares also the optical heart rate monitor of our system with a commercial pulse oximeter device.



Fig. 6. Test Management Sub-System view: The medical staff can see the images visualized on the screen of the patient, augmented with position of the gaze marked with a red circle and the heart rate (in this case, the patient is watching the middle of the cup).

information [32]

$$C_{\text{PPG}}^i(t) = \frac{\sum_{x,y \in \Omega_{\text{ROI}}} C^i(x,y,t)}{|\Omega_{\text{ROI}}|}, \quad i \in \{R, G, B\} \quad (1)$$

where $C^i(x,y,t)$ is the pixel value at the coordinates (x,y) in the i channel at the time t .

This feature works thanks to the photoplethysmography (PPG) technique, which can capture optical absorption characteristics of (oxy-)hemoglobin [32]. In short, when the heart pumps, the average optical intensity in the ROI changes. This change can be caught by a standard webcam. By processing this information, we can estimate the patient's heart rate and then the heart rate variability (HRV) that is the fluctuation of the length of heart beat intervals. HRV is closely related to emotional arousal and stress [33].

In Fig. 5 we show the visual interface presented by the "Test Management Sub-System" during the calibration procedure. We noticed the face of the observer enclosed in a green ellipse, as he is well positioned. The BPM are also visualized; for sake of clarity we have added a pulse oximeter to check and compare the BPM evolution.

D. Single-Board Computer Version

Using a single-board computer, such as the Raspberry PI4 [34] for implementing the "Measurement Sub-System", we encountered significant limitations in the processing speed. To overcome these limitations, we included in our software the possibility that the "Measurement Sub-System" locally records the frames from the camera while the observer watches the pairs of images. When the test is over, the "Measurement Sub-System" transfers the complete data gathered by the camera to the "Test Management Sub-System", which calculates and save in a database all the predictions on each frame and the corresponding heart rate.

E. Sub-systems synchrony

As already stated, the proposed eye-tracking system is composed by the "Measurement Sub-System" aimed to interact with the patient and the "Test Management Sub-System" used by medical staff for perform the calibration,

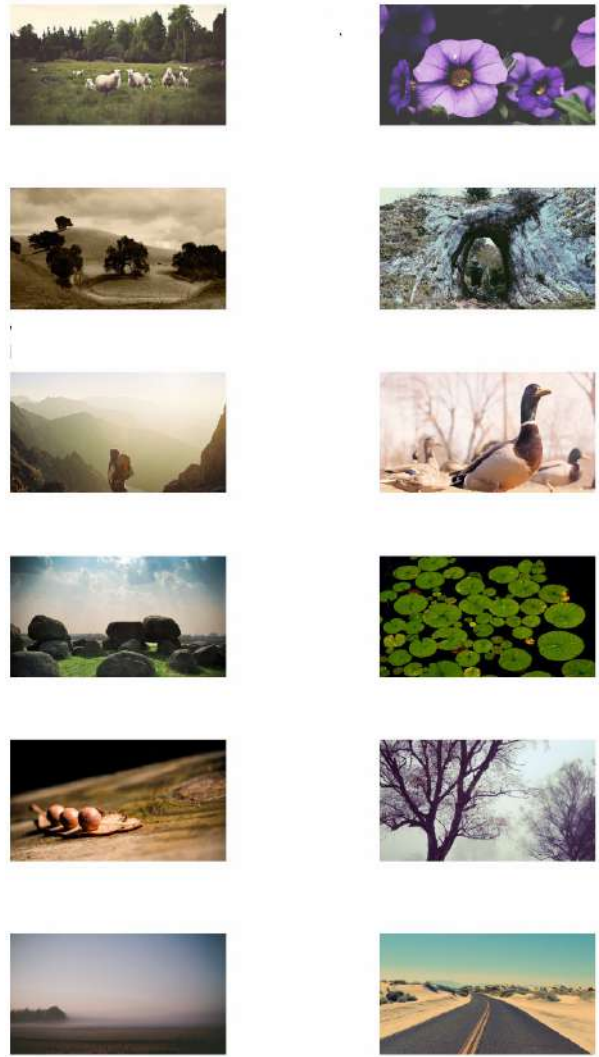


Fig. 7. MCI Visual Paired Comparison test: Images used in the familiarization phase.

configuring the test protocol, recording the patient data, monitoring the test and storing the test results.

The two subsystems are connected to each other in a sub-net created via a router or an access point and the "Test Management Sub-System" can also be connected to the network of the hospital.

During the test, the same pair of images visualized on the screen of the "Measurement Sub-System" will appear in synchrony on the screen of the "Test Management Sub-System", augmented with the position of the gaze marked with a red circle (see Fig. 6) and the heart rate.

III. EXAMPLE OF TEST PROTOCOL

In collaboration with the physicians of the Neurological Clinic of the University of Genoa, we are defining the details of a protocol for the implementation of a test for the early diagnosis of Mild Cognitive Impairment based

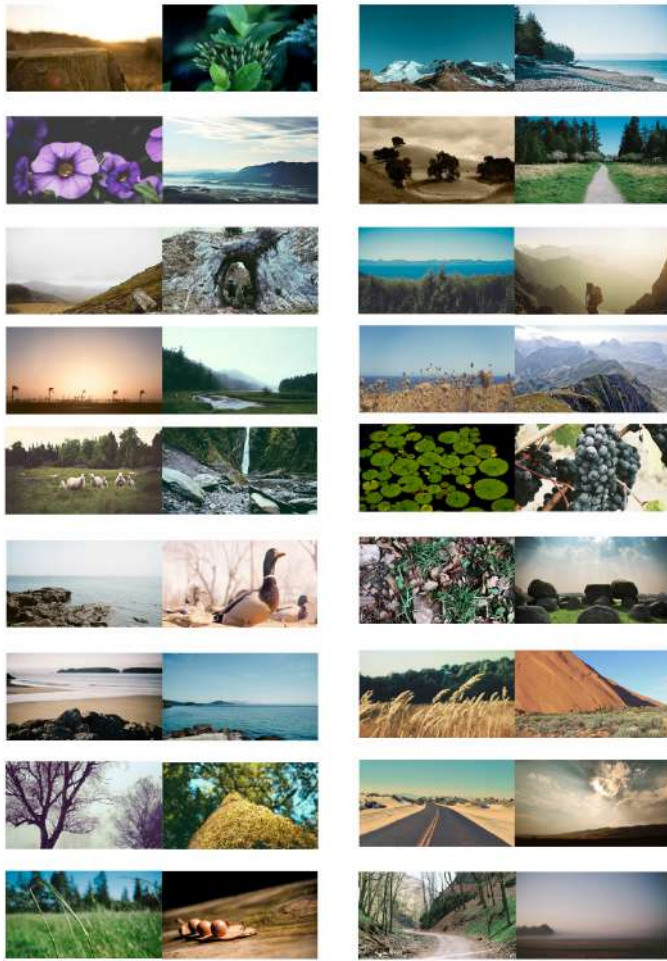


Fig. 8. MCI Visual Paired Comparison test: Examples of pair of images displayed in the test phase.

on Visual Paired Comparison, with the aim of testing it in clinics with geriatric patients.

A possible protocol employs 36 images, of which 12 are presented in the familiarization phase, and 24 in the testing phase only

During the familiarization phase (total duration 3 minutes) the images are shown on the display for 15 seconds each, accompanying it with a specific question aimed at reinforcing its memorization, e.g., "What kind of animal is shown?" (see Fig. 7). After the familiarization phase, the test phase starts (see Fig. 8). In this phase (total duration 4 minutes) we show in random sequence:

- six pairs of new images;
- six pairs with a new image on the left and a known image from the familiarization phase on the right;
- six pairs with a new image on the right and a known image from the familiarization phase on the left.

As already discussed, we expect that healthy people will observe the new stimuli about 70 % of the time, while observers with MCI will not give preference to either of the two stimuli [14].

IV. DISCUSSION AND CONCLUSIONS

In this paper we have presented a system for eye-tracking aimed at carrying out tests based on a Visual Paired Comparison protocol for the early detection of Mild Cognitive Impairment. The proposed system uses webcams, standard personal computers that may already be available at hospital structures (or even low cost-single board computers) and exploits some efficient machine learning algorithms.

Even if the proposed system is not the first affordable eye-tracker that uses artificial intelligence algorithm, standard webcams and personal computers (see, e.g., [21], [35]), it is characterized by its specialization and validation for tests based on Visual Paired Comparison protocol and for its integration with stress measurement based on heart rate variability. Given the low-cost of the proposed eye-tracking system, once the measurement protocol has been defined and its clinical trials concluded, a wide diffusion of the test is expected, with great positive repercussions on public health, given that the Alzheimer's disease is one of the chronic diseases that affects a large segment of the elderly population and is spreading more and more also in low-income countries, with the rise in life expectancy.

REFERENCES

- [1] Jennifer Romano Bergstrom and Andrew Schall. *Eye Tracking in User Experience Design*. Morgan Kaufmann, 2014.
- [2] David A. Robinson. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions on Bio-medical Electronics*, 10(4):137–145, 1963.
- [3] Deouell LY, Keren AS, Yuval-Greenberg S. Saccadic spike potentials in gamma-band eeg: characterization, detection and suppression. *Neuroimage*, 49(3):2248–63, 2010.
- [4] K. Schreiber and T. Haslwanter. Improving calibration of 3-d video oculography systems. *IEEE Transactions on Biomedical Engineering*, 51(4):676–679, 2004.
- [5] Mark M. J. Houben, Janine Goumans, and Johannes van der Steen. Recording Three-Dimensional Eye Movements: Scleral Search Coils versus Video Oculography. *Investigative Ophthalmology & Visual Science*, 47(1):179–187, 01 2006.
- [6] F. Z. Javaid, J. Brenton, L. Guo, and M. F. Cordeiro. Visual and Ocular Manifestations of Alzheimer's Disease and Their Use as Biomarkers for Diagnosis and Progression. *Frontiers in neurology*, 7:55, 01 2016.
- [7] Dmitry Lagun, Cecelia Manzanares, Stuart M. Zola, Elizabeth A. Buffalo, and Eugene Agichtein. Detecting cognitive impairment by eye movement analysis using automatic classification algorithms. *Journal of Neuroscience Methods*, 201(1):196–203, 2011.
- [8] RC Petersen, GE Smith, SC Waring, RJ Ivnik, EG Tangalos, and E Kokmen. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol.*, 56(3):303–308, 1999.
- [9] Serge Gauthier et al. Mild cognitive impairment. *The Lancet*, 367(9518):1262–1270, 2006.
- [10] C. M. Clark, C. Davatzikos, A. Borthakur, A. Newberg, S. Leight, and J. Q. Lee, V. M. andTrojanowski. Biomarkers for early detection of alzheimer pathology. *Neuro-Signals*, 16(1):11–18, 2008.
- [11] Stuart M Zola, CM Manzanares, P Clopton, JJ Lah, and AI Levey. A behavioral task predicts conversion to mild cognitive impairment and alzheimer's disease. *American Journal of Alzheimer's Disease & Other Dementias*, 28(2):179–184, 2013.

- [12] Trevor J Crawford, Steve Higham, Ted Renvoize, Julie Patel, Mark Dale, Anur Suriya, and Sue Tetley. Inhibitory control of saccadic eye movements and cognitive impairment in alzheimer's disease. *Biological psychiatry*, 57(9):1052–1060, 2005.
- [13] Julius Opwonya, Changwon Wang, Kyoung-Mi Jang, Kunho Lee, Joong Il Kim, and Jaeuk U. Kim. Inhibitory control of saccadic eye movements and cognitive impairment in mild cognitive impairment. *Frontiers in Aging Neuroscience*, 14, 2022.
- [14] Michael D. Crutcher, Rose Calhoun-Haney, Cecelia M. Manzanares, James J. Lah, Allan I. Levey, and Stuart M. Zola. Eye tracking during a visual paired comparison task as a predictor of early dementia. *American Journal of Alzheimer's Disease & Other Dementias*, 24(3):258–266, 2009. PMID: 19246573.
- [15] JR Manns, CE Stark, and LR Squire. The visual paired-comparison task as a measure of declarative memory. *Proc Natl Acad Sci U S A*, 97(22):12375–12379, 2000.
- [16] SM Zola, LR Squire, E Teng, L Stefanacci, EA Buffalo, and RE Clark. Impaired recognition memory in monkeys after damage limited to the hippocampal region. *J Neurosci.*, 20(1):451–463, 2000.
- [17] A Nuthmann. Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychon Bull Rev.*, 24(2):370–392, 2017.
- [18] Guido Van Rossum et al. Python programming language. In *USENIX annual technical conference*, volume 41, pages 1–36, 2007.
- [19] Paul Viola and Michael J Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.
- [20] Rainer Lienhart and Jochen Maydt. An extended set of Haar-like features for rapid object detection. In *Proceedings. international conference on image processing*, volume 1, pages I–I. IEEE, 2002.
- [21] Raimondas Zemblys, Diederick C Niehorster, Oleg Komogortsev, and Kenneth Holmqvist. Using machine learning to detect events in eye-tracking data. *Behavior research methods*, 50(1):160–181, 2018.
- [22] A. Haar. On the theory of orthogonal function systems. *Mathematische Annalen*, 69:331—371, 1910.
- [23] Franklin Crow. Summed-area tables for texture mapping. In *SIGGRAPH '84: Proc 11th annual conference on Computer graphics and interactive techniques*, pages 207–212, 1984.
- [24] J.P. Lewis. Fast template matching. *Proc. Vision Interface*, pages 120–123, 1995.
- [25] R.E. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. *Machine Learning*, 37:297–336, 1999.
- [26] Ivan Culjak, David Abram, Tomislav Pribanic, Hrvoje Dzapo, and Mario Cifrek. A brief introduction to opencv. In *2012 proceedings of the 35th international convention MIPRO*, pages 1725–1730. IEEE, 2012.
- [27] PyQT. Pyqt reference guide. 2012.
- [28] Martín Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [29] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. The MIT Press, 2016.
- [30] Marian Stewart Bartlett. Real time face detection and facial expression recognition: Development and applications to human computer interaction. In *2003 Conference on Computer Vision and Pattern Recognition Workshop*, volume 5, pages 53–53, 2003.
- [31] Isayiyas Nigatu Tiba and Lei Li. Image-based automatic pulse rate monitoring system using pc webcam. *International journal of engineering research and technology*, 2, 2013.
- [32] Wim Verkruyssen, Lars O Svaasand, and J Stuart Nelson. Remote plethysmographic imaging using ambient light. *Opt. Express*, 16(26):21434–21445, Dec 2008.
- [33] HG Kim, EJ Cheon, DS Bai, YH Lee, and BH Koo. Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investig.*, 15:235–245, 2018.
- [34] Raspberry PI foundation.
- [35] F. Corno, L. Farinetti, and I. Signorile. A cost-effective solution for eye-gaze assistive technology. In *Proceedings. IEEE International Conference on Multimedia and Expo*, volume 2, pages 433–436 vol.2, 2002.