Facial expression recognition from nao robot within a memory training program for individuals with mild cognitive impairment

1st Berardina De Carolis
Department of Computer Science
University of Bari "A. Moro"
Bari, Italy
berardina.decarolis@uniba.it

2nd Giuseppe Palestra
Research Department
HERO srl
Apulia, Italy
giuseppepalestra@gmail.com

3rd Olimpia Pino
Department of Psychology
University of Parma
Parma, Italy
olimpia.pino@unipr.it

Abstract—Mild Cognitive Impairment refers to a borderline state between healthy aging and dementia. Memory-training program plays a crucial role in the reduction of the possible conversion in dementia and a robot mediated memory training is useful to overcome limits of traditional programs. The present study addresses the effectiveness of a system in automatically recognize facial expression from video recorded sessions of a robot mediated memory training lasted 2 months involving 21 patients. The system is able to recognize facial expressions from group sessions handling partially occluded faces. Findings showed that in all participants the system is able to recognize facial expressions.

Index Terms—facial expression recognition, social robot, memory training, Mild Cognitive Impairment

I. Introduction

Mild Cognitive Impairment (MCI) concerns a stage between normal aging and early dementia marked by cognitive deficit characterized by scores below the norm on psychometric tests, preserved functional abilities and high levels of quality of life [1]. The prevalence of MCI in individuals >65 years of age is between 10 and 20%. MCI is highly likely to convert in dementia at a rate of about 13% per year and in the rest of the patients the impairment persists stable or even return to normal over time [2]. Dementia was estimated to have been detected around the globe at the rate of one new patient about every 7 seconds. Therefore, MCI has become a relevant research topic because it could play a critical role in distinguishing developmental changes in lifespan memory from those that are real signs of the disorder. Delaying the onset of dementia by as little as one year could decrease the global burden of Alzheimer's by 9 million of patients in 2050 [3]. In order to maintaining cognitive functions non-pharmacological programs are developed. These programs involve qualified psychologist and therapists in order to conduct new tasks and new exercises, to monitor the performance of the patients, to

SAT19: 1st Workshop on Socio-Affective Technologies: an interdisciplinary approach, October 7, 2019, Bari, Italy

provide helpful feedback, to analyze patients' performances over the time [4]. The psychologists' primary objective for MCI patients is to keep their cognitive ability when functional capabilities and independence are not compromised [2]. The worldwide incidence of MCI expects to increase in the next few years [3]; however, space and personnel shortages are already becoming a problem owing to an unprecedented rise in life expectancy [3]. In very recent years, new tools and technologies based on machine learning and robotics are successfully applied to the field of psychology and could be used to in memory training program for people with MCI. Humanoid robots are able to improve mood, emotional expressiveness and social relationships among patients with dementia [5]– [7] also executing many assistive functionalities [8]–[10] and providing life assistance demonstrating that the information support provided by the robot also has the potential to improve the daily life of persons with a mild level of dementia [5]. Most recent advances in information and communication technologies have enabled the development of telepresence robots to connect a family member and a person with dementia as a means of enhancing communication between these two parties [11]. The humanoids skills are progressively enhanced: they are able to recognize faces, call people by their name, shape their behavior considering the mood of people interacting with them [12], [13]. Some robots can also reproduce emotions [12], [14], making their human mate feel welcomed, and simulating empathy [15], [16]. Kinetics technology can help them reproduce movements [17], while speech recognition software allows them to respond to what people say, even in many different languages. During human-robot interaction, the mirror-circuit, responsible for social interaction, is verified to be active [18] suggesting that humans can consider robots as real companions with their own intentions. Many studies have employed the robot NAO. If appropriately programmed, it is able to decode human emotions, simulate emotions through the color of his eyes or the position of the body, recognize faces and model physical exercise to a group of seniors [19], and equipped to measure of health and environmental parameters [20]. Robotics could partially fill in some of the identified gaps in current health care and home care/self-care provisions for promising applications in these fields that we expect to play relevant roles in the near future. With emerging research suggesting that mobile robot systems can improve elderly care [6], [7], [10], also through the development of a coding system aimed at measuring engagement-related behavior across activities in people with dementia [21], [22]. With the growing incidence of pathologies and cognitive impairment associated with aging, there will be an increasing demand for maintaining care systems and services for elderly with the imperative of economic cost-effectiveness of care provision. In our previous work [23], NAO has been evaluated as mediator in a memorytraining program for people with MCI in a center for cognitive disorders therapy.

The focus of the present paper was evaluate the feasibility and usefulness of NAO platform in cognitive stimulation. In human interaction emotional and social signals expressing additional information are essential. For this aim, particular attention was paid to quantifying the effectiveness of robot on well being of the training recipients and this was realized measuring facial expressions at the same time gaze when human participants interact with the synthetic agent. This study aims to exploit video clips recorded during the 2-month experiment of the previous study analyzing facial expressions of the participants in the experiment. The system presented is able: i) to recognize group facial expression thanks a multiface detector; ii) to recognize facial expression in partially occluded faces. This paper is organized as follows. Section 2 presents Materials and Methods used. Section 3 reports experimental results. Finally, conclusions are drawn.

II. Materials and Methods

A. Participants

The participants were selected from the population of outpatients attending the Center for Cognitive Disorders and Dementia of AUSL Parma (Italy). All the participants were firstly evaluated by memory-disorders specialists. The diagnosis of MCI was based on a detailed medical history, relevant physical and neurological examinations, negative laboratory findings, and neuroimaging studies. Subjects are enrolled according to the following inclusion criteria: a) diagnosis of MCI obtained through Petersen guidelines, and full marks in the two tests measuring daily living activities (ADL and IADL); b) both genders; c) chronological age comprised between 45 and 85 years; and d) without pharmacological treatment. Exclusion criteria were a diagnosis of major neurocognitive disorder (defined using DSM 5 criteria), history of symptomatic stroke (although silent brain infarction was not an exclusion), history of other central nervous system diseases, serious medical or psychiatric illness that would interfere with study participation, such Parkinson’s disease, HIV/AIDS, or other contraindications. Informed consent was obtained from all the patients or from their legal representatives when appropriate. 21 individuals (10 females and 11 males) participated in the experiment with a mean age of 73.45 years (SD = 7.71). The mean education level was of 9.90 years (SD = 4.58) with a minimum value of 5 years corresponding to the conclusion of the elementary school and a maximum value of 18 years which corresponds to a bachelor degree.

B. Robot Mediated Memory Training Program

The memory training exercises were implemented on NAO on the basis of exercises described in literature [24] and aimed to train: i) focused attention; ii) divided and alternate attention; and iii) categorization and association as learning strategies. Five tasks were implemented in NAO, considering the characteristics of the robot:

1) Reading stories;
2) Questions about the story;
3) Associated/not associated words;
4) Associated/not associated word recall;
5) Song-singer match;

C. Video corpus

In this study, we used a corpus of 48 memory program session video clips of one hour from 24 therapeutic sessions. Each video clips recorded three or four participants. These videos were recorder during two months by two cameras placed in the therapeutic room and they have been used to record all participants. Overall, at the end of the experiment for each participant a total of eight hours of video have been collected.

D. Group Facial Expression Recognition System

The group facial expression recognition system detect faces in the corpus and then recognize 6 basic emotions (anger, disgust, fear, happiness, sadness and surprise) plus neutral expression. The video analysis system is based on a previous study that aimed to recognize six basic emotions through facial expression [25] but the system has been improved in order: i) to work in groups using a multi-face detector; ii) handle partial occlusions of the face. In group facial expression recognition handle occlusions is very important in order to ensure a high accuracy rate. Indeed, sometimes the face of the participants is occluded by a hand as well as by an arm of the other participants as shown in Figure 1.

Fig. 1. Examples of occlusion in the video corpus.

After identifying a face, the system extracts, for each face, facial landmarks locating 77 key points. Once the 77 points are identified, the software tracks linear, polygonal, elliptical and angular characteristics, i.e. the distance between two points to find the following: three lines describing the left eyebrow; two defining the left eye; one for the cheeks; one for nose; eight for the mouth. The system then determines polygonal features, calculating the area delimited by irregular polygons created using three or more key reference points, specifically: one for the left eye; one forming a triangle between the corners of the left eye and the left corner of the mouth; one for the mouth. Thus, the system traces the elliptic characteristics, calculated by the ratio between the major axis and the minor axis of the ellipse, in particular seven ellipses are chosen between the reference points: one for the left eyebrow; three for the eye, left upper and lower eyelid; three for the mouth, lower and upper lips. The pipeline of the system is depicted in Figure 2. The system has been tested on the Extended Cohn-Kanade (CK+) data set, a well known facial expression image database of 123 individuals of different gender, ethnicity and age. The system reach an average facial expression recognition accuracy of 95,46% on six basic facial expression and an average accuracy of 94,24% on 6 basic facial expression plus neutral.

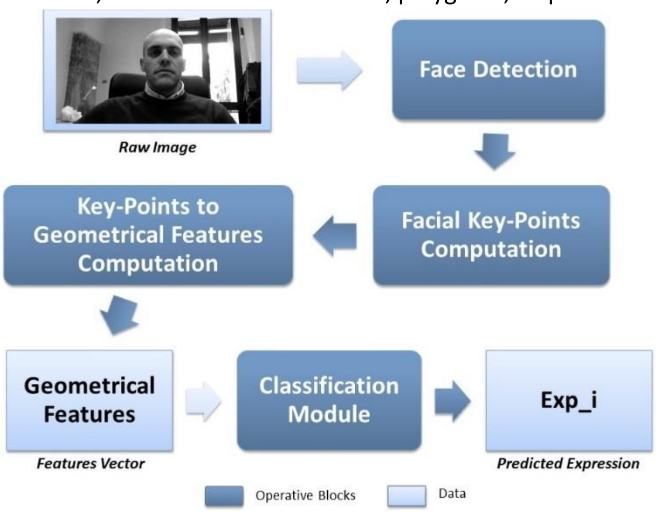

Fig. 2. Facial expression analyzing software flow diagram.

Overall, the system extracts 32 geometric features that have been used in total or in part (to handle occlusions) in order to train a model. To recognize facial expression, the system uses a classification module that, through a Random Forest classifier, analyzes the geometric characteristic vectors to determine the facial expression.

III. Results

In order to test the system, the video clips of the corpus have been down sampled at 1 frame per second. For each participant 28,800 frames have been analyzed (1 frame/second x 8 hours of recording). To evaluate the facial expression recognition during the memory training program mediated by the robot: i) the number of detected face (nFD); ii) and the number of each facial expression recognized (nFE) for each frame for each participants in the video corpus have been used as metrics. Overall, the system was able to analyze all the video corpus and all the 21 faces of the participants have been detected. With respect to nFD, it has been observed that, in percentage, a face has been detected in total of the corpus with a success rate of 56%. Moreover, respect to the nFE, it has been observed that in each frame where a face had been detected (also if partially occluded) the system was able to recognize a facial expression. In percentage, the three most common facial expressions are neutral that has been founded in the 41% of the frames, happiness in the 17% and sadness in the 15% as shown in Figure 3.

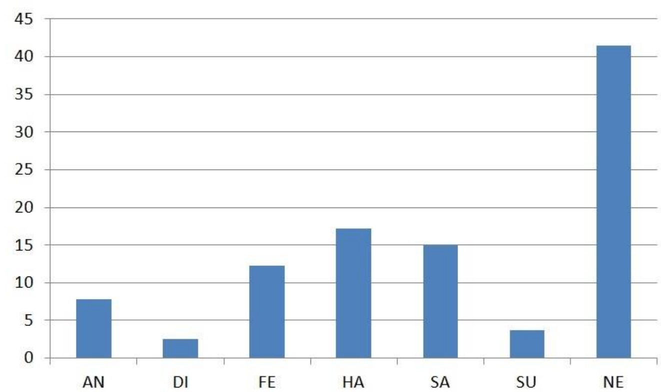

Fig. 3. Facial expression recognized in the video corpus in percentage.

In Table I are reported the number of frames where a face had been detected and Table II reports the number of facial expression recognized for each participant divided into facial expressions.

TABLE I

Number of Face Detected (nFD) for each participant (#).

Participant (#)	nFD
1	23,567
2	22,678
3	16,709
4	18,541
5	20,341
6	16,354
7	15,980
8	12,434
9	11,345
10	16,709
11	18,541
12	13,678
13	12,454
14	13,544
15	25,123
16	16,341
17	11,490
18	10,431
19	14,235
20	12,357
21	18,235

IV. Conclusions

In this paper a system for group facial expression recognition with handling of partial occlusions has been presented as well as its application to the field of psychology. A particular application such as a memory training program for MCI in adults can benefit from computer vision and machine learning technologies to understand better how patients react to robot mediated memory training program. Moreover, from automatic facial expression recognition psychologists can keep track of the mood of the individuals involved during the training sessions. In future work, other analysis on the video corpus will be done to understand the engagement, how the participants have been involved during the memory training program and how to better involve the individuals with MCI in new memory training program mediated by a social robot.

TABLE II

Number of Facial Expression recognized (nFE) for each participant (#) divided into facial expressions (AN=anger, DI=disgust, FE=fear, HA=happiness, SA=sadness, SU=surprise, NE=neutral).

#	nFE	AN	DI	FE	HA	SA	SU	NE
1	23567	987	345	2312	4054	2345	870	12654
2	22678	1022	1233	1982	3281	1757	765	12638
3	16709	998	374	1534	2918	3982	551	6352
4	18541	1543	235	2234	3002	4526	549	6452
5	20341	738	397	2391	3384	1902	847	10682
6	16354	543	235	1534	3002	1138	549	9353
7	15980	637	332	932	3363	1332	321	9063
8	12434	1101	356	2291	2308	2347	394	3637
9	11345	1438	234	1982	1221	1395	254	4821
10	16709	1412	316	1535	2237	4223	332	6654
11	18541	873	339	1982	3922	2023	748	8654
12	13678	667	254	1625	3323	2145	782	4882
13	12454	823	443	1912	3231	1198	256	4591
14	13544	2383	487	1112	1005	1345	870	6342
15	25123	1587	429	2542	4124	2891	794	12756
16	16341	673	345	1612	2905	945	651	9210
17	11490	2012	234	1234	1123	3487	870	2530
18	10431	970	221	873	3130	1881	703	2653
19	14235	2189	884	3823	1290	3321	218	2510
20	12357	1123	543	2912	2931	2283	214	2351
21	18235	987	345	1982	2052	2345	870	9654

References

[1] O. Pino, "Memory impairments and rehabilitation: Evidence-based effects of approaches and training programs," *The Open Rehabilitation Journal*, vol. 8, no. 1, 2015.

[2] S. Gauthier, B. Reisberg, M. Zaudig, R. C. Petersen, K. Ritchie, K. Broich, S. Belleville, H. Brodaty, D. Bennett, H. Chertkow *et al.*, "Mild cognitive impairment," *The Lancet*, vol. 367, no. 9518, pp. 1262– 1270, 2006.

[3] T. G. Ton, T. DeLeire, S. G. May, N. Hou, M. G. Tebeka, E. Chen, and J. Chodosh, "The financial burden and health care utilization patterns associated with amnestic mild cognitive impairment," *Alzheimer's & Dementia*, vol. 13, no. 3, pp. 217–224, 2017.

[4] O. Pino, *Ricucire i ricordi: la memoria, i suoi disturbi, le evidenze di efficacia dei trattamenti riabilitativi*. Mondadori universita, 2017.`

[5] S. Sabanovic, C. C. Bennett, W.-L. Chang, and L. Huber, "Paro robot affects diverse interaction modalities in group sensory therapy for older adults with dementia," in *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–6.

[6] H.-M. Gross, C. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, A. Bley, T. Langner, C. Martin, and M. Merten, "I'll keep an eye on you: Home robot companion for elderly people with cognitive impairment," in *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2481–2488.

[7] F. Mart´ın, C. Aguero, J. M. Ca¨ nas, G. Abella, R. Ben˜ ´ıtez, S. Rivero, M. Valenti, and P. Mart´ınez-Mart´ın, "Robots in therapy for dementia patients," *Journal of Physical Agents*, vol. 7, no. 1, pp. 48–55, 2013.

[8] C. D. Kidd and C. Breazeal, "Robots at home: Understanding longterm human-robot interaction," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE, 2008, pp. 3230–3235.

[9] B. Graf, C. Parlitz, and M. Hagele, "Robotic home assistant care-o-bot®¨ 3 product vision and innovation platform," in *International Conference on Human-Computer Interaction*. Springer, 2009, pp. 312–320.

[10] H.-M. Gross, C. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, A. Bley, T. Langner, M. Merten, C. Huijnen, H. van den Heuvel *et al.*, "Further progress towards a home robot companion for people with mild cognitive impairment," in *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 637–644.

[11] W. Moyle, C. Jones, M. Cooke, S. ODwyer, B. Sung, and S. Drummond, "Connecting the person with dementia and family: a feasibility study of a telepresence robot," *BMC geriatrics*, vol. 14, no. 1, p. 7, 2014.

[12] B. De Carolis, S. Ferilli, and G. Palestra, "Simulating empathic behavior in a social assistive robot," *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 5073–5094, 2017.

[13] B. De Carolis, N. Macchiarulo, and G. Palestra, "Soft biometrics for social adaptive robots," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 2019, pp. 687–699.

[14] B. N. De Carolis, S. Ferilli, G. Palestra, and V. Carofiglio, "Towards an empathic social robot for ambient assisted living." in *ESSEM@ AAMAS*, pp. 19–34. [Online]. Available: https://dl.acm.org/citation.cfm?id=3054107

[15] V. Gallese, C. Keysers, and G. Rizzolatti, "A unifying view of the basis of social cognition," *Trends in cognitive sciences*, vol. 8, no. 9, pp. 396– 403, 2004.

[16] B. De Carolis, S. Ferilli, and G. Palestra, "Improving speech-based human robot interaction with emotion recognition," in *International*

Symposium on Methodologies for Intelligent Systems. Springer, 2015, pp. 273–279.

[17] B. S. Ertugrul, C. Gurpinar, H. Kivrak, and H. Kose, “Gesture recogni-ˇ tion for humanoid assisted interactive sign language tutoring,” in *Signal Processing and Communications Applications Conference (SIU), 2013 21st*. IEEE, 2013, pp. 1–4.

[18] V. Gazzola, G. Rizzolatti, B. Wicker, and C. Keysers, “The anthropomorphic brain: the mirror neuron system responds to human and robotic actions,” *Neuroimage*, vol. 35, no. 4, pp. 1674–1684, 2007.

[19] I. Back, K. Makela, and J. Kallio, “Robot-guided exercise program for¨ the rehabilitation of older nursing home residents,” *Annals of Long Term Care*, vol. 21, no. 6, 2013.

[20] J. P. Vital, M. S. Couceiro, N. M. Rodrigues, C. M. Figueiredo, and N. M. Ferreira, “Fostering the nao platform as an elderly care robot,” in *2013 IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, 2013, pp. 1–5.

[21] G. Perugia, R. van Berkel, M. D´ıaz-Boladeras, A. Catala-Mallofr` e,´ M. Rauterberg, and E. Barakova, “Understanding engagement in dementia through behavior. the ethographic and laban-inspired coding system of engagement (elicse) and the evidence-based model of engagementrelated behavior (emodeb),” *Frontiers in psychology*, vol. 9, 2018.

[22] P. Boissy, H. Corriveau, F. Michaud, D. Labonte, and M.-P. Royer,´ “A qualitative study of in-home robotic telepresence for home care of community-living elderly subjects,” *Journal of telemedicine and telecare*, vol. 13, no. 2, pp. 79–84, 2007.

[23] O. Pino, G. Palestra, R. Trevino, and B. De Carolis, “The humanoid robot nao as trainer in a memory program for elderly people with mild cognitive impairment,” *International Journal of Social Robotics*, pp. 1–13, 2019.

[24] D. Gollin, A. Ferrari, and A. Peruzzi, *Una palestra per la mente. Stimolazione cognitiva per l’invecchiamento cerebrale e le demenze*. Edizioni Erickson, 2007.

[25] G. Palestra, A. Pettinicchio, M. Del Coco, P. Carcagni, M. Leo, and C. Distante, “Improved performance in facial expression recognition using 32 geometric features,” in *International Conference on Image Analysis and Processing*. Springer, 2015, pp. 518–528.