Modern Environmental Science and Engineering (ISSN 2333-2581)
November 2020, Volume 6, No. 11, pp. 1213-1217
Doi: 10.15341/mese(2333-2581)/11.06.2020/008
Academic Star Publishing Company, 2020
www.academicstar.us

Facial Feature Analysis for Pseudodementia: A Preliminary Study

Brian Sumali[1], Yasue Mitsukura[1], and Taishiro Kishimoto[2]

1. Department of System & Design Engineering, Faculty of Science and Technology, Keio University, Japan

2. Department of Neuropsychiatry, School of Medicine, Keio University, Japan

Abstract: Pseudodementia is a type of temporary cognitive impairment caused by mental health disorders, differing from true dementia with root cause of neurological disorders. The most common cause is depression, and the comorbidity of dementia with depression in elderly patients deludes even expert psychologists. Although pseudodementia can be diagnosed with extensive testing, it is time consuming and taxing for both the psychiatrists and the patient, as at least two tests must be taken — one for depression screening and another for dementia screening. Additionally, although machine learning has been utilized in automated mental health screening, attention for pseudodementia is minimal, with progression only in discussing and proposing pseudodementia tests in the medical field. In this research we aim to extract facial features from actual dementia patients and depression patients, which were conducted and diagnosed by licensed clinical psychiatrists. The features extracted from dementia patients and depression patients served as the basis for screening pseudodementia. We also tried to utilize machine learning using the facial features to examine the possibility of automated pseudodementia screening. As it is a preliminary study, a conventional machine learning model was utilized along with popular feature selection algorithm. Satisfactory result of 81.7% accuracy was obtained, although improvement must be made for actual clinical implementation. Several features that were chosen by the machine learning was also reported, which may be beneficial for human clinician.

Key words: pseudodementia, facial features, machine learning

1. Introduction

Dementia is commonly associated with irreversible, declining cognitive function: memory problems (both short-term and long-term), apparent confusion, reduced concentration, and apathy. However, sometimes other factor may cause the apparent cognitive loss and depending on the cause, some cases are reversible. One of such factors are depression.

Depression cases where the patient also affected with apparent loss of cognitive function is defined as “Pseudodementia”. This is especially different from other causes of reversible cognitive loss, because the patients only appear to be affected with cognitive decline and comprehensive evaluation of those patient reveal normal cognitive abilities. Pseudo-dementia is difficult to diagnose, even for expert psychologists and to this date, some depression cases still are mislabeled as dementia cases, making treatment ineffective and even dangerous to the patients [1].

Conventional psychoanalysis assessment studies utilized facial features (gaze, blink, emotion detection, etc.) [2, 3], biosignals such as electroencephalogram, heart rate, respiration [4-7], and auditory features (intensity, tone, speed of speech, etc.) [8, 9]. Although biosignals are the most reliable data source, most of the measurement devices are arduous to equip, limiting their value. In the other hand, facial and acoustic features may be obtained with minimal burden to the patient.

Corresponding author: Yasue Mitsukura, Professor; research areas/interests: bio-signal processing, medical engineering. E-mail: mitsukura@keio.jp.

The objective of this paper is to investigate the possibility of differentiating dementia patients and depression patients based on face feature analysis. Our hypothesis was, if classification of dementia patient and depression patient is possible, then machine learning might be able to be utilized to detect pseudodementia as the next step. This study is a part of The Project for Objective Measures Using Computational Psychiatry Technology (PROMPT) [10]. The rest of this paper is organized as follows. In second part, the data acquisition protocol is described. In third part, the analysis of obtained data is explained. The results are shown in fourth part and the discussed in the fifth part. The paper is then concluded in the sixth part.

2. Data Acquisition

The data was obtained by recording a patient's clinical interview session with a therapist. A full clinical session consisted of 10 minutes free talk session followed by 20 to 30 minutes rating session.

In the free talk session, the therapist conducts a typical interview concerning the patient's daily life and mood. Although the results of this session do not contribute for the assessment of the patient, the interview still follows guidelines and is a semi-structured interview. The length of this segment is around 10 minutes.

In the rating session, the patient is interviewed based on a clinical assessment tools related to their mental health history, which may include some additional tasks and tests such as clock-drawing test and memory test for dementia screening or some personal questions such as their sleep habit and depressive mood in the recent weeks which are related to depression screening. The duration of a single rating segment typically lasts more than 20 minutes.

During the interview, the patient and the therapist were seated across a table. On the therapist's side a laptop for controlling the recording system and a microphone facing the therapist were set. On the patient's side, a microphone and a video camera, both facing the patient, were placed. Fig. 1 shows the recording environment. The therapist started the recording before the session began and ended it after the session finished. The experiment was conducted on Keio University Hospital and Joint Medical Research Institute. The experiment was approved by Keio University Hospital Ethics Committee (20160156, 20150427).

3. Analysis

The flow-chart utilized during the analysis for classification using machine learning was depicted in Fig. 2. An explanation for each part of the flowchart is given in the following sub-sections. To prevent the model learning age feature instead of disease's features and to increase the contrast between the features, we screened the dataset and only include datasets which satisfy the following criteria:

(1) Age between 57 and 84 years-old. This is to remove the effect of aging, which is known to be positively correlated with dementia symptoms

(2) For dementia patients: mini-mental state examination (MMSE) score of 24 or less accompanied with geriatric depression scale (GDS) score of 4 or less; This is to ensure that the dementia patients are symptomatic and are not afflicted with depression co-morbidity.

(3) For depression patients: 17-item Hamilton depression rating scale (HAMD17) of 8 or more; Similar with dementia patient criteria, this is to ensure that the depression patients are symptomatic. Depression patients that have co-morbidities with dementia are always classified as dementia patient.

Fig. 1 Recording setup.

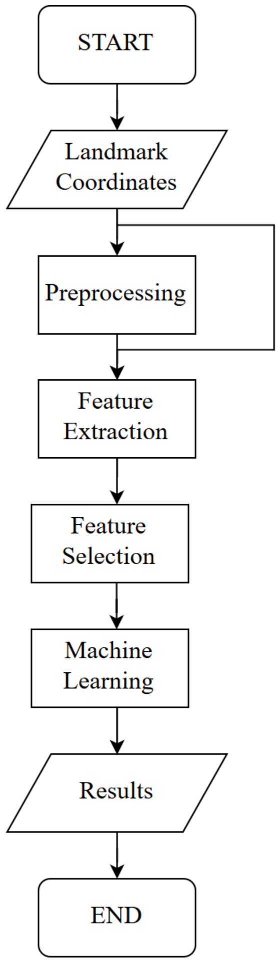

Fig. 2 Flowchart of data analysis.

The qualifying dataset were 75 datasets from 57 subjects, consisting of 46 dementia datasets and 29 depression datasets. To protect the identity of the subjects, facial landmarks extraction was performed using Omron's OKAO Vision [11] to extract the X-Y coordinates, as seen in Fig. 3. These 40 points of facial landmarks are processed and analyzed instead of the raw face images.

3.1 Preprocessing

The obtained facial landmarks were then normalized, such that the center of face lies on origin coordinates (0,0) and then each landmark coordinate set was divided according to the face width and height: X-coordinates were divided by face width and Y-coordinates were divided by face height as shown on Eqs. (1) and (2).

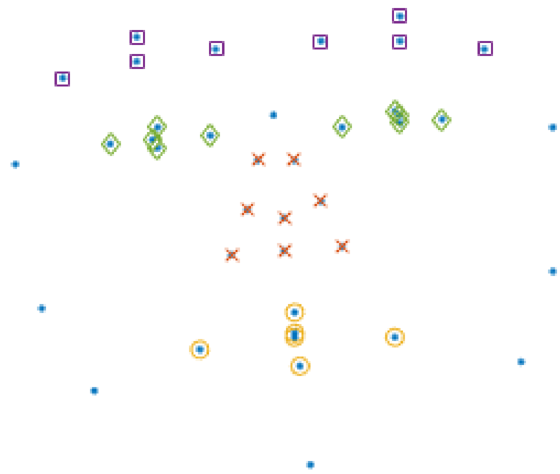

Fig. 3 Extracted facial landmarks.

$$\widehat{X_i} = \frac{(X_i - X_c)}{Width} \tag{1}$$

$$\widehat{Y_i} = \frac{(Y_i - Y_c)}{Height} \tag{2}$$

Another preprocessing step was performed to remove outliers, by removing frames which landmarks have value less than 1st percentile and greater than 99th percentile.

3.2 Feature extraction

After the preprocessing, the feature extraction was performed. Features extracted in this study were: speed statistics of each landmark, speed statistics of the face center, area of mouth, and the standard deviation of eye pupils' position. The speed of each landmark was computed using the Eqs. (3) and 4.

$$S_i = \sqrt{\left(\hat{X}_i - \hat{X}_{i+1}\right)^2 + \left(\hat{Y}_i - \hat{Y}_{i+1}\right)^2} \tag{3}$$

$$\widehat{S_j} = \sum_{i=30(j-1)+1}^{30j} S_i \tag{4}$$

where $\hat{X}_i$ and $\hat{Y}_i$ denotes a preprocessed landmark's coordinate on frame i. Here, S_i is the landmark i's *speed per frame* and $\widehat{S_j}$ denotes the landmark j's *speed per second*. The constant 30 represents the camera's frame-per-second (FPS) rate of 30.

The area of left eye, right eye, and mouth were computed using the Eq. (5), which is a general equation of computing the area of an arbitrary polygon in 2D space.

$$A = \frac{1}{2}\left|\sum_{i=0}^{n-1} \hat{X}_i Y_{i+1} - \hat{X}_{i+1} Y_i\right| \quad (5)$$

when $i = n - 1$, then $i + 1$ is expressed as 0.

3.3 Feature Selection

Feature selection is an important process to reduce the number of features used as a machine learning input. Several important reasons to perform feature selection are to mitigate overfitting problem, to make the model easier to interpret, and to improve the speed of a model's learning session. Here, we utilized Least Absolute Shrinkage and Selection Operator (LASSO) algorithm [12] as the feature selection algorithm. The top 10% features that is frequently selected with this algorithm are utilized for machine learning.

3.4 Machine Learning

We then examined the possibility of differentiating depression patient and dementia patient based on facial features by utilizing supervised machine learning with 10-fold cross validation to measure the model's performance. The machine learning model we used was support vector machines (SVM) with various kernels: linear, polynomial with order of 2, and radial basis function (RBF).

4. Results

Twenty-two (22) features were selected with the LASSO algorithm. These features were utilized for machine learning. The selected features are listed on Table 1.

The best performances of the SVM models were the one with polynomial kernel with order of 2. Its average accuracy is 81.37±15.03%. The second best was SVM with Linear kernel, with average accuracy of 77.86±16.54% whilst RBF kernel performed the worst, with the average accuracy of 69.46±15.96%. This confirms our hypothesis of classifying depression versus dementia patients.

Table 1 Features selected by LASSO algorithm.

Statistical Feature	Landmark
Average speed	Left pupil, left eye (bottom), right pupil, left nose, right nose, right eyebrow (bottom & left), right jaw, right ear (bottom)
Median speed	Right pupil, upper lip (bottom)
Standard deviation of speed	Left eye (top), right eye (top & bottom), glabella, left eyebrow (bottom), right eyebrow (top, bottom, and left), left jaw, right jaw, left ear (bottom), right ear (top)
95th percentile of speed	Right nose, left eyebrow (top & left), right eyebrow (bottom). left ear (bottom)

5. Conclusion and Future Work

Our results show that the dementia patients and depression can be clearly classified even with traditional machine learning technique such as SVM. This result suggests the possibility of automatic pseudodementia screening by utilizing machine learning.

Our future work would be to utilize data from dementia patients with co-affliction of depression and to test other machine learning models.

Acknowledgement

As a part of PROMPT project, this research was funded by the Japan Agency for Medical Research and Development (AMED) under Grant Number JP18he1102004. The Grant was awarded on 29 October 2015 and ended on 31 March 2019.

References

[1] A. S. Kelley, K. McGarry, R. Gorges and J. S.Skinner, The burden of health care costs for patients with dementia in the last 5 years of life, *Ann. Intern. Med*. (2015).

[2] J. Li, Z. Liu, Z. Ding and G. Wang, A novel study for MDD detection through task-elicited facial cues, in: *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018.

[3] Z. S. Shah, K. Sidorov and D. Marshall, Psychomotor cues for depression screening, in: *Int Conf Digit Signal Process DSP*, 2017

[4] G. Henderson, E. Ifeachor, N. Hudson, C. Goh, N. Outram and S. Wimalaratna et al., Development and assessment of methods for detecting dementia using the

human electroencephalogram, *IEEE Trans Biomed Eng*. (2006).

[5] S. Khatun, B. I. Morshed and G. M. Bidelman, Single channel EEG based score generation to monitor the severity and progression of mild cognitive impairment, in: *2018 IEEE Int Conf Electro/Information Technol*, 2018.

[6] R. Nakamura and Y. Mitsukura, Feature analysis of electroencephalography in patients with depression, in: *2018 IEEE Life Sciences Conference (LSC)*, 2018.

[7] H. Uchida, H. Yamaguchi, S. Kobashi, Y. Hata, N. Tsuchiya and H. Nakajima, Fuzzy estimation system of dementia severity using biological information during sleep, *IEEE Int Conf Fuzzy Syst*, 2009.

[8] N. Linz, J. Tröger, J. Alexandersson, M. Wolters, A. König and P. Robert, Predicting dementia screening and staging scores from semantic verbal fluency performance, in: *IEEE Int. Conf. Data Min Work ICDMW*, 2017.

[9] S. Harati, A. Crowell, H. Mayberg and S. Nemati, Depression severity classification from speech emotion, in: *Conf Proc. Annu Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2018.

[10] T. Kishimoto et al., The project for objective measures using computational psychiatry technology (PROMPT): Rationale, design, and methodology, *BMC Psychiatry* (2019).

[11] S. Lao and M. Kawade, Vision-based face understanding technologies and their applications, *Advances in Biometric Person Authentication* (2004).

[12] V. Fonti and E. Belitser, Feature selection using LASSO, 2017.