March-April 2020
ISSN: 0193-4120 Page No. 7749 - 7753

Emotion Classification Based on EEG using Independent Component Analysis and Genetic Algorithm

[1] Bimo Rian Tri Nugroho, [2]Sugondo Hadiyoso, [3] Rita Purnama Sari, [4] Inung Wijayanto
[1,3,4] School of Electrical Engineering, [2]Telkom Applied Science School, Telkom University Bandung, Indonesia
[1]bimoriantrinugroho@gmail.com, [2]sugondo@telkomuniniversity.ac.id, [3]ritapurnama@telkomuniversity.ac.id,
[4]inungwijayanto@telkomuniversity.ac.id.

Article Info
Volume 83
Page Number: 7749 - 7753
Publication Issue:
March - April 2020

Article History
Article Received: *24 July 2019*
Revised: *12 September 2019*
Accepted: *15 February 2020*
Publication: *09 April 2020*

Abstract

Emotion is a psychological condition as a reaction to an event. All forms of emotions are thought to be controlled by the central nervous system, the brain. A human's emotions are often associated with changes in facial expressions as markers. Several studies have successfully carried out emotional detection through facial imaging. But the detection of emotions through these expressions tends to be hidden. Therefore, in this research study emotions through EEG signals to overcome these problems because EEG characteristics are unique and difficult to hide. Emotional EEG signals come from the DEAP database in the form of arousal, valence, liking, dominance, and familiarity with video as the stimulus. The raw EEG signal processed by Independent Component Analysis (ICA) to obtain features is then classified by Support Vector Machine (SVM) combined with Genetic Algorithm (GA) for optimization. The simulation carried out gave an accuracy of 77.27% on the classification of sad and happy emotions.

Keywords*; *Emotion, detection, EEG, ICA, SVM, GA

I. INTRODUCTION

reaction to an event. Emotion is one of the complex biological processes controlled by the brain [1]. Emotions can affect health conditions in terms of one's body metabolism and psychology. Especially negative emotions that must be avoided because they can have bad consequences. Emotions can influence ways of thinking, making decisions and solving problems [2]. This is what underlies many studies to analyze, detect or classify emotions. Some methods are done, namely psychological tests, characterization through facial expressions and what is now attracting much attention is the study of electroencephalogram (EEG) related to emotions.

Emotional recognition through EEG characterization is used in medical applications and scientific research [1]. The objectives are various including knowing the sources of emotional generation, control or biofeedback for psychological health and some applied to the brain computer interface. Referring to this goal, in this paper, we have a high interest in analyzing EEG waves for emotional classification even though this study still encountered many difficulties, especially the recording of data, validation and the limitations of the EEG related emotion database. As initial research, in this paper we use open databases sourced from the Database for Emotion Analysis using Physiological Signals (DEAP) for characterization and EEG analysis related to emotion.

DEAP is a database that contains a person logistical research. Emotions are given about participants' facial expressions and brain wave results. Research was carried out using various stimuli such as

Published by: The Mattingley Publishing Co., Inc.

images, sounds, or videos [3]. The stimulus is given to all participants, so they could know which participants were attending. The parameters are given by the brain when expressing participants are happy, sad, or neutral. This change in condition affects the EEG wave and gives characterization to each emotional condition. Several studies have proposed the use of various methods for emotional classification based on DEAP dataset [1], [4], [5]. The method of analysis in the time domain, frequency, time-frequency was applied to obtain the EEG feature. The feature reduction method is also applied to computational efficiency. The machine learning algorithm is applied to classification and the value of accuracy is a measure of the performance of the proposed method. However, the average accuracy achieved has not been satisfactory. Thus, efforts to develop methods that are accurate and resource efficient in classification refer to the data set method. Therefore in this paper, a method for classifying emotions is proposed in the EEG DEAP dataset.

This study involved two human emotion groups that represented happy and sad based on the valence scale. This study uses Independent Component Analysis (ICA) for EEG feature extraction and Support Vector Machine (SVM) as classifiers. Test scenarios where parameters are optimized by Genetic Algorithm (GA) are also conducted to evaluate the impact of GA on accuracy. Finally, the results of this research are expected to be a reference for the development and application of a larger database in the future.

II. MATERIAL AND METHOD

A. EEG Dataset

Data on this research is sourced from the DEAP research test [6]. This dataset is taken from 32 subjects ranging in age from 19-37 (14 women, 18 men). Each participant recorded 32 EEG channels while watching a music video. Video is used as a stimulus to arouse the emotions of participants. The EEG sampling frequency is 256 Hz with a high-pass filtering to 2 Hz cut-off frequency. The experiment was conducted 5 times to see the level of valence with the video given. The following Table 1 contains details of the subject's condition including the level of valence.

Table 1. Subject's profile

Subject	Age	Sex	Valence	Class
1	31	Male	3.92	Sad
2	24	Female	5,01	Happy
3	19	Female	4,37	Happy
4	24	Female	8,05	Happy
5	24	Male	6,96	Happy
6	23	Male	3,49	Sad
7	31	Male	2,05	Sad
8	22	Female	5,79	Happy
9	25	Female	4,04	Happy
10	31	Female	7,14	Happy
11	27	Female	8,7	Happy
12	37	Male	8,21	Happy
13	24	Female	9	Happy
14	27	Female	3,99	Sad
15	22	Female	4,05	Happy
16	28	Male	3,64	Sad
17	25	Male	6,05	Happy
18	29	Male	4,23	Happy
19	27	Male	1,62	Sad
20	25	Male	6,35	Happy
21	30	Male	4,83	Happy
22	28	Female	6,03	Happy
23	27	Male	5,09	Happy
24	28	Female	7,03	Happy
25	26	Female	3,05	Sad
26	36	Male	5,57	Happy
27	35	Male	6.08	Happy
28	24	Male	5,27	Happy
29	24	Male	5,09	Happy
30	33	Male	5,33	Happy
31	21	Female	5,17	Happy
32	33	Female	5,31	Happy

B. Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a method to find information from a set of data where independent factors are traced, to separate independent signals mixed and the signal will be recorded by several sensors [7]. In the mixing model, s is the signal source before it is mixed with other signals, A is the mixing matrix, and x is mixed signals. The ICA output in this research is used as a feature set.

$$x = As \quad (1)$$

where s :

$$s = A^{-1}x \quad (2)$$

C. Support Vector Machine (SVM)

Support Vector Machine is a technique for predicting a case such as classification and has a linear classifier principle and has been developed to solve problems for non-linear classifier [8]. In the classification of SVM it is usually used in the classification of two classes or multiclass. SVM works with a kernel approach that has a function to map low dimensions to higher dimensions so that the results achieved reach the optimal level. The kernel functions used are polynomial, gaussian, and linear kernels. Parameters other than the kernel are scale kernels whose scale values are 1-10 and polynomial orders whose scale values are 1-10.

D. Genetic Algorithm (GA)

Genetic algorithms are studies that study evolutionary mechanics and methods to determine the minimum value of a parameter and to find feature optimization and selection. The parameter space formed by genetic algorithms can include multilayer perceptions of each layer. Genetic algorithms use techniques from evolutionary biology such as mutations, inheritance and crossover [9].

The genetic algorithm method is used to obtain the optimum solution from many possible solutions [9]. In a study that uses more possibilities, exact methods are usually not used even though genetic algorithms do not always provide optimal solutions in optimization problems but through processes, genetic algorithms can get good solutions [9].

In the GA research, it will classify or classify data according to the conditions we want. Before that there is training first in order to get the parameters that are suitable for the desired conditions and the testing of the parameters specified.

In applying the GA algorithm to feature-selection problems, each chromosome is represented by a dimension m binary vector, where m is the total number of features. If it's a little 1, then the related feature is included, and if it's a little 0, the feature is not included. The GA process for feature selection problems is the same as GA. This process ends when he finds a feature subset with the highest accuracy or reaches the maximum number of iterations [10]. In this research, GA is proposed to find the best features so that it can improve the performance of the classifier.

In the system kernel functions, kernel scale, and polynomial orders from Support Vector Machine (SVM) are parameters that are optimized by Genetic Algorithm (GA). GA looks for the optimum solution for each system process. Each iteration has the possibility of different results so it takes as many as 20 iteration experiments so that conclusions can be drawn which optimum solution is obtained. There are process of genetic algorithms are as follows [11]:

1. Determine the chromosome number, generation, and mutation rate and cross rate value
2. Generate population chromosomes, and the values of initial chromosome gene chromosomes with random values
3. Process steps 4-7 until the number of generations is met
4. Evaluate the chromosome fitness value by calculating the objective function
5. Chromosome selection
6. Crossover
7. Mutations
8. Solution (Best Chromosome)

Published by: The Mattingley Publishing Co., Inc.

TEST Engineering & Management

March-April 2020
ISSN: 0193-4120 Page No. 7749 - 7753

III. RESULTS AND DISCUSSION

The signal features of ICA results are obtained in the form of a matrix. The matrix contains the attribute features measuring [a b] each channel electrode. The features obtained are A is the mixing matrix and W is the separation matrix. Table 2 below is an example of the ICA feature from one of the electrodes.

Table 2. ICA feature

No	*Mixing*	*Separating*	Class
1	-3.88	-0.25	Sad
2	35.46	-0.02	Happy
3	-5.85	-0.17	Happy
4	-70.25	-0.01	Happy
5	-16.82	-0.05	Happy
6	8.19	0.12	Sad
7	-12.25	-0.08	Sad
8	-7.01	-0.14	Happy
9	-14.12	-0.07	Happy
10	22.43	0.04	Happy
11	-26.8838	-0.03	Happy
12	16.55	0.06	Happy
13	-9.04	-0.11	Happy
14	10.90	0.09	Happy
15	-8.99	-0.11	Happy
16	-15.16	-0.06	Happy
17	-13.30	-0.07	Happy
18	-13.56	-0.07	Happy
19	-53.31	-0.01	Happy
20	-5.30	-0.18	Happy
21	8.62	0.11	Sad
22	84.59	0.01	Happy
23	-7.71	-0.12	Happy
24	-52.38	-0.01	Happy
25	-10.44	-0.09	Sad
26	18.78	0.05	Sad
27	-6.59	-0.15	Happy
28	-18.88	-0.05	Sad
29	-11.00	-0.09	Happy
30	20.42	0.04	Happy
31	-10.54	-0.09	Happy
32	-10.52	-0.09	Happy

The system performance parameter is how many classes correctly detected are expressed in accuracy. In this study, EEG datasets were divided into two groups, namely training data and test data. We apply the ratio of test data and training data to 0.7 / 0.3; 0.5 / 0.5 and 0.3 / 0.7.

To get optimum performance, the parameters of SVM are optimized with GA. The first functions are kernel functions, namely polynomial, gaussian and linear, the second scale of the kernel which has a scale of 1-10 and third scale is a polynomial order that has a scale of 1-10 so GA looks for the optimum solution of 300 possibilities that might occur in the system. These parameters become input for GA in getting the optimum solution. Based on the test results, the system is accurate with a ratio of 0.7 / 0.3 with an accuracy value of 77.27% (Table 3).

Table 3. Comparison of the results of system accuracy

Scenario (Test/training data)	Best Accuracy	
	SVM	SVM optimized GA
0.7/0.3	56.25%	77.27%
0.5/0.5	56.25%	75%
0.3/0.7	50%	60%

System testing proves that the classification results that only use SVM have smaller accuracy values than the SVM classification results optimized by GA. The best accuracy results are 0.7 / 0.3 with a value of 77.27%. The results of the classification of proposed methods in this research are shown in Table 4.

Table 4. Classification results

No	System output 1	System output	Emotion classification
1	Happy	Happy	True
2	Happy	Happy	True
3	Happy	Happy	True
4	Happy	Happy	True
5	Sad	Happy	False

Published by: The Mattingley Publishing Co., Inc.

TEST Engineering & Management

March-April 2020
ISSN: 0193-4120 Page No. 7749 - 7753

6	Sad	sad	True
7	Happy	Happy	True
8	Happy	Happy	True
9	Happy	Happy	True
10	Happy	Sad	False
11	Happy	Happy	True
12	Happy	Happy	True
13	Sad	Sad	True
14	Happy	Happy	True
15	Happy	Happy	True
16	Sad	Happy	False
17	Happy	Happy	True
18	Happy	Happy	True
19	Happy	Happy	True
20	Happy	Happy	True
21	Sad	Happy	False
22	Sad	Happy	False

IV. CONCLUSION

In this research a system for the classification of human emotions has been simulated based on EEG signals. The dataset is source from DEAP. Two groups of data namely happy and sad refer to the valence scale, have been simulated. Feature extraction using ICA is applied to get a number of attributes as distinguishing characters for each group. Genetic Algorithm is also applied, aiming to reduce features and determine the best features so as to improve classification performance. GA proved to be able to increase accuracy compared to SVM without GA optimization. This research resulted in good accuracy of 77.27% with the comparison of test data and training data 0.7 and 0.3 respectively. The results of this research provide the knowledge that ICA is quite competitive for feature extraction referring to the small number of features. In addition, GA is very effective for use in the purpose of improving the performance of classifiers.

REFERENCES

[1] J. Liu, H. Meng, A. Nandi, and M. Li, "Emotion detection from EEG recordings," *2016 12th Int. Conf. Nat. Comput. Fuzzy Syst. Knowl. Discov. ICNC-FSKD 2016*, pp. 1722–1727, 2016.

[2] N. Jung, C. Wranke, K. Hamburger, and M. Knauff, "How emotions affect logical reasoning: evidence from experiments with mood-manipulated participants, spider phobics, and people with exam anxiety," *Front. Psychol.*, vol. 5, pp. 1–34, 2014.

[3] U. of L. Queen Mary, "DEAP: A Dataset for Emotion Analysis using Physiological and Audiovisual Signals," vol. 3, no. 1, pp. 18–31.

[4] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion Recognition from EEG Signals Using Multidimensional Information in EMD Domain," *Biomed Res. Int.*, vol. 2017, pp. 1–9, 2017.

[5] J. Liu, H. Meng, M. Li, F. Zhang, R. Qin, and A. K. Nandi, "Emotion detection from EEG recordings based on supervised and unsupervised dimension reduction," *Concurr. Comput.* , vol. 30, no. 23, pp. 1–13, 2018.

[6] S. Koelstra, C. Muhl, M. Soleymani et al., "DEAP: a database ¨ for emotion analysis; using physiological signals," IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18–31, 2012.

[7] A. Hyvärinen and E. Oja, "Independent Component Analysis: Algorithms and Applications," *Neural Networks*, vol. 13, no. 45, pp. 411–430, 2000.

[8] A. Rizal and S. Hadiyoso, "Sample Entropy on Multidistance Signal Level Difference for Epileptic EEG Classification," *Sci. World J.*, vol. 2018, pp. 1–6, 2018.

[9] Anies Hannawati, Thiang, and Eleazar, "Pencarian Rute Optimum Menggunakan Algoritma Genetika," *J. Tek. Elektro*, vol. 2, no. 2, pp. 78–83, 2002.

[10] B. Nakisa, M. N. Rastgoo, D. Tjondronegoro, and V. Chandran, "Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors," *Expert Syst. Appl.*, vol. 93, no. October, pp. 143–155, 2018.

[11] M. Gen and R. Cheng, "Genetic algorithms and engineering design," *John Wily Sons, New York*, 1997.