The performance of fuzzy clustering for the bimodal stress recognition

Qiang Li

Research School of Computer Science
Australian National University
Email: u6999647@anu.edu.au

Abstract. There are many people experiencing stress regularly. Chronic stress results in disease and negatively affects people's psychological health. Thus, developing a robust method for the rapid and accurate detection of human stress is of great importance. In this paper, we apply the fuzzy clustering technique to classify the stress level, using ten features from ANUStressDB which contains images of both RGB and thermal modalities. The results show that a two-layer neural network is about 3% better performance than the cascade networks. The fuzzy c-means clustering has a lower performance than the two-layer neural network, with 50.48% vs 64.52%.

Keywords: Fuzzy clustering, deep Learning, thermal video, Constrcasc neural network, neural network

1 Introduction

Nowadays, stress recognition is of great importance as stress affects people's health conditions. A robust with rapid detection of stress may enable the continuous monitoring of stress and facilitate people to accept effective treatment for stress-related illness. Frequently used measurements involve traditional methods such as physical symptoms (movement of lips, mouth, and eyebrows [4], however, more and more vision-based techniques such as RGB video recorder or thermal imaging have been employed for stress recognition [2]. Using computer vision techniques helps to avoid the contact with the user and then increase the quality of data collection and further the accuracy of model prediction.

There are many approaches to analysis for the vision-based dataset. Past research has primarily used traditional machine learning approaches, such as decision trees, support vector machine, k-nearest neighbor and others [7]. Recently, Neural network usually is selected for the classification tasks, although simple feedforward neural network still faces generalization and convergence speed challenges. Because of this, some updated algorithms have been proposed, such as Constrcasc which has a good time efficiency and generalization ability. This paper aims to compare the performance of the above three mentioned methods' performance on the stress recognition by using the vision-based dataset. The three methods include simple neural network, Constrcasc neural network and fuzzy c-means clustering.

Part 2 describes the fuzzy clustering and Constrcasc neural network algorithm, the dataset, and the training processes. Part 3 provides the results and discussion. Part 4 brings up the future works.

2 Method

2.1 Fuzzy c-mean approach

The first Fuzzy sets theory is provided by Zadeh in 1996 [6] which was used to deal with problems where human is lack of knowledge. Fuzzy clustering is one of the most widely used method based on the fuzzy logic theory. Fuzzy c-mean (FCM) is one of the clustering techniques proposed by Bezdek [8] based on the iterative optimization. It is different from the hard c-means clustering that the membership degree of data in a FCM can have a value in a range of [0,1]. Given a dependent feature vector x, FCM aims to minimize the objective function in the followings:

$$A = \sum_{i=1}^{n} \sum_{k=1}^{c} u_{ik * \|x_i - c_k\|^2}^{m}$$

Where A is the objective function, c is the number of clusters, u_{ik}^{m} is the degree of membership of x_i in cluster k, x_i is the ith value of D-dimensional measured data, c_k is the D-dimension center of the kth cluster. The distance between x_i and c_k, $\|x_i - c_k\|^2$ is calculated as the Euclidean distance. The degree membership u_{ik}^{m} can be calculated as:

$$u_{ik} = \left(\frac{1}{\sum_{l=1}^{c} \frac{\|x_i - c_k\|}{\|x_i - c_i\|}}\right)^{\frac{2}{m-1}}$$

The process of FCM clustering could be summarized in the following steps:

1. Determine the number of classes (c).
2. Calculate the centers of clusters.
3. For each point, compute the membership value for each cluster. The value of the weight parameter (m) defines the amount of fuzziness of the clustering process.
4. Compute the objective function. The optimization process in the FCM algorithm can continue up to r iterations.
5. Assign each positioning index to a cluster after defuzzification.

2.2 ConstrCasc Neural Network

ConstrCasc algorithm is a modification from the cascade correlation algorithm provided from [1]. In the paper, it is used for the face recognition. Compared with the traditional cascade correlation algorithm, ConstrCasc reduces the complexity of the network but retain the performance. The procedure of the algorithm is as followings:

1. Begin with a minimal network which contains one output layer.
2. Freeze the neurons (except the output layer), train the model until the stop condition satisfied.
3. If the loss is good, stop training process.
4. If not, freeze the weights of the current network.
5. Add a new hidden layer, and train the network based on optimizing the correlation between the output of the hidden layer and the last residual error.
6. Go to the Step 2.

In this paper, some changes have been made based on the characteristics of the dataset.
First, a simple neural network substitutes the convolutional neural network (CNN) as used in the paper. The format of the thermal video dataset is group of super-pixels which are adjacent pixels having similar characteristic and special information (color/temperature). Because the dataset is on the type of image vectors, the two-dimensional layer in the paper is changed to one dimensional layer here.
Second, different from the original paper that a 16 by 16 hidden layer, there are only two cascade layers with each layer adding one hidden neuron, 10 neurons in the input layer and 2 output neurons for the two class labels. Like cascade correlation algorithm, the neuron is added to network during training process. The new hidden neuron is connected to the input neurons and fully connected to the output neurons. The algorithm freezes the weights of previous trained network.

2.3 Dataset

The thermal video dataset comes from the database-ANUStressDB for stress recognition that contains images of both RGB and thermal modalities. The data covers 31 subjects with a blend of sex and ages.
The RGB and thermal modalities were captured by a FLIR infrared camera and a Microsoft webcam. Both cameras were working at 30 frame/second at a 640*480 pixels resolution. The tester watched the film with a collection of positive or negative clips as stress stimulator.
There are 5 seconds blank screen between the clips for giving the moment to tester to neutralize their emotions. At the end of the experiment, the participants will fill a questionnaire survey for assigning the frames as “stressful/calm”. The dataset is post-processed to represent a thermal I mage as a group of super-pixels. A super-pixel is a group of adjacent pixels with similar color (temperature). The paper uses Linear Spectral Clustering (LSC) method to compute super-pixels (see figure 1). More detailed thermal detection methodology and super-pixel extracting technique could be seen in [2].

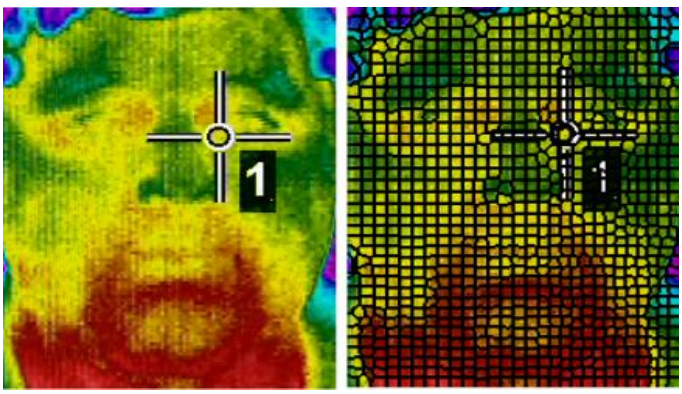

Figure 1. Super-pixel method on facial region [2]

2.4 Training Methodology

For classifying the features, a z-score method has been adopted as a measure to standardize the range of the different features. The z-score enables comparison of data of different magnitudes.

Three experimental scenarios are provided: 1) two-layer neural network, 2) two cascade layer and 3) fuzzy c-mean.

Two-layer neural network
For each 5-fold split, 5000 epochs are run. During the experiment, the accuracy of prediction is volatile for 500 and 2000 epochs. The learning algorithm applied is Adam for all learning process and loss is computed by the CrossEntropyLoss for every epoch. Leaky-relu is used as the activation function. Different learning rate has been tested and finally it is set up to 0.01.

ConstrCasc
In the cascade network, the activation function selected is the hyperbolic tangent function, with the Resilient propagation (RPROP) as the learning algorithm as proposed in the reference paper [1]. The RPROP algorithm has the advantages of faster convergence speed [5]. The cross-entropy loss is selected because the probability of labels is calculated. The data are normalized before the training. A 5-fold cross-validation is conducted and repeated for 5000 epochs. All the above process are repeated for adding one cascade layer and adding two cascade layers.

Fuzzy c-mean clustering
The maximum iteration is 500. We have tested different fuzzy parameter and use 2 as the fuzzy parameter because it produces the best results. The number of clusters is 2 (stress/calm).

3 Results

The average performance of three different network is shown in Table 1.

Table 1. The average performance of each network

Network	**Two-layer neural network**	**Two cascade layers**	**Fuzzy c-means**
Average of testing accuracy (%)	64.52	61.29	50.48

Based on the average performance of all three networks in Table 1, a two-layer neural network is about 3% better performance than the cascade networks. The small sample of testing sets might limit the generalization of the ConstrCasc models. However, the ConstrCasc networks are more time efficiency and less violate on the testing accuracy. The fuzzy c-means clustering has a lower performance than the two-layer neural network, with 50.48% vs 64.52%. Comparing to the mean accuracy rate of 62% in the thermal paper [2], this paper has a slightly higher performance.

4 Conclusion and future work

In this paper, we make the performance comparison of the simple neural network, the ConstrCasc network and the fuzzy c-mean clustering method by using thermal and RGB modalities and temperature of super-pixels to decide if a user is in a stress situation or not. In terms of the test results, the fuzzy c-mean clustering method performed worse than the other two deep learning methods. The accuracy rate for the fuzzy c-mean clustering method is 50.48%, which is about 14% lower than the neural network, and 11% lower than the cascade layer network result. In terms of the time efficiency, the fuzzy c-means has a higher time efficiency than the other two methods. The results demonstrate that for the small sample of testing dataset like this paper, a simple layer neural network might be enough for the accurate classification.
There are some limitations in this paper which can be further evaluated in future. The dataset contains few data points with only 620*11 entries which is a small sample. A larger and more representative dataset could be tested in future to achieve more accurate analysis. Cascade correlation network has all the previously added hidden units frozen which could lead to poor detection of pre-existing hidden units. A CasPer algorithm could be conducted to address the problem. Fuzzy c-means algorithm is sensitive to the outliers and high dimensional data sets. A more variety of fuzzifier parameter could be tested in the future.

References

1. Khoo, S., & Gedeon, T. (2008, November). Generalisation Performance vs. Architecture Variations in Constructive Cascade Networks. In International Conference on Neural Information Processing (pp. 236-243). Springer, Berlin, Heidelberg.
2. Irani, R., Nasrollahi, K., Dhall, A., Moeslund, T. B., & Gedeon, T. (2016, December). Thermal super-pixels for bimodal stress recognition. In 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1-6). IEEE.
3. Sharma, N., Dhall, A., Gedeon, T., & Goecke, R. (2014). Thermal spatio-temporal data for stress recognition. EURASIP Journal on Image and Video Processing, 2014(1), 1-12.
4. Metaxas, D., Venkataraman, S., & Vogler, C. (2004, June). Image-based stress recognition using a model-based dynamic face tracking system. In International Conference on Computational Science (pp. 813-821). Springer, Berlin, Heidelberg.
5. Riedmiller, M., & Braun, H. (1993, March). A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In IEEE international conference on neural networks (pp. 586-591). IEEE.
6. Zadeh LA (1996) Fuzzy sets. In: Zadeh LA (ed) Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers. World Scientific, Singapore, pp 394–432.
7. Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on intelligent transportation systems, 6(2), 156-166.
8. Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. Computers & geosciences, 10(2-3), 191-203.