Biologically Inspired Cognitive Architectures xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Biologically Inspired Cognitive Architectures

journal homepage: www.elsevier.com/locate/bica

Research article

Image-based emotion recognition using evolutionary algorithms

Hadjer Boubenna*, Dohoon Lee

Department of Electrical and Computer Engineering, Pusan National University, Republic of Korea

ARTICLE INFO

Keywords:
Pattern recognition
Evolutionary algorithms
Optimization
Machine learning
Supervised learning

ABSTRACT

In pattern recognition, the classification accuracy has a strong correlation with the selected features. Therefore, in the present paper, we applied an evolutionary algorithm in combination with linear discriminant analysis (LDA) to enhance the feature selection in a static image-based facial expressions system. The accuracy of the classification depends on whether the features are well representing the expression or not. Therefore the optimization of the selected features will automatically improve the classification accuracy. The proposed method not only improves the classification but also reduces the dimensionality of features. Our approach outperforms linear-based dimensionality reduction algorithms and other existing genetic-based feature selection algorithms. Further, we compare our approach with VGG (Visual Geometry Group)-face convolutional neural network (CNN), according to the experimental results, the overall accuracy is 98.67% for either our approach or VGG-face. However, the proposed method outperforms CNN in terms of training time and features size. The proposed method proves that it is able to achieve high accuracy by using far fewer features than CNN and within a reasonable training time.

Introduction

Emotion recognition based on static images is one of the hardest pattern recognition problems, because facial expressions can be ambiguous and may have several possible interpretations. Thus for vision community, developing an automatic facial expression recognition system is important. Facial expression recognition has increasingly been used in different fields such as entertainment, military, medicine, E- learning, monitoring and marketing.

The big size of features is a crucial issue for facial emotion recognition system and other pattern recognition systems. When the search space is very large a curse of dimensionality problem (Gheyas & Smith, 2010) occurs which will drop the classification accuracy.

Thus optimizing the selection of features is indispensable to overcome this problem and improve the classification accuracy. Feature selection has been widely applied to many machine learning applications, such as classification, regression and clustering (Dash & Liu, 1997). Features selection techniques can be grouped into four categories which are filter (Roffo & Melzi, 2016; Roffo, Melzi, & Cristani, 2015; Zhang et al., 2015), hybrid (Huerta, Duval, & Hao, 2006), wrapper (Chuang, Yang, & Yang, 2009; Zhang, Wang, Phillips, & Ji, 2014) and embedded (Duval, Hao, & Hernandez Hernandez, 2009; Hernandez, Duval, & Hao, 2007) based features selection.

The first category is called wrapper, this method is based on a predictive model, where each new subset is used for training and tested on the hold-out set.

And then the score is calculated based on the error of the model (for the given subset). Wrapper methods are computationally expensive however they provide a good performance (Kohavi & John, 1997). Instead of using a predictive model, filter based methods use a proxy measure such as mutual information (Kohavi & John, 1997) and pointwise mutual information (Guyon & Elisseeff, 2003). Filter based methods are computationally less expensive than wrappers based methods however they provide lower performance than wrappers (Kohavi & John, 1997). The third category is called embedded, this method performs the selection of features as part of the model construction process. Embedded techniques are computationally more complex than filters and less complex than wrappers techniques (Kohavi & John, 1997). A hybrid feature selection can be a combination of filters and wrappers. First, filters are applied to select the features, and then wrappers are applied to further refine the features (Yang & Pedersen, 1997).

Finding an optimal subset of features in a reasonable time (Hsu, Hsieh, & Lu, 2011; Liu & Motoda, 2007) is a hard task, however an approximation of the subset of features can be a solution to overcome the time complexity. Therefore we suggest involving an optimization technique in the task of feature selection to reduce the search space and improve the classification accuracy. Our objective is to build an automatic emotion recognition system using bio-inspired optimization to decrease the redundancy, and keep only the relevant features that

* Corresponding author.
E-mail addresses: hadjer@pusan.ac.kr (H. Boubenna), dohoon@pusan.ac.kr (D. Lee).

https://doi.org/10.1016/j.bica.2018.04.008
Received 29 March 2018; Received in revised form 14 April 2018; Accepted 15 April 2018

encode important information.

We propose a genetic algorithm (GA) in combination with linear discriminant classifier (LDA) to evaluate each combination of features, we run the algorithm for a finite number of iterations until converging to the optimum subset of features which will later be used to classify the emotion.

The rest of the paper is structured as follows: in Section "Related Work" we describe the related works. In Section "Framework" we give an overview of the proposed framework. In Section "Experiments and analysis" we describe the experimental results. In Section "Conclusion" we give the conclusion.

Related work

Feature selection is indispensable in any pattern recognition system, in which the dimensionality of feature vectors will be decreased significantly. Decreasing the redundant features will increase the classification accuracy automatically (Janecek, Gansterer, Demel, & Ecker, 2008).

Principal component analysis (PCA) is considered as one of the most encountered feature dimensionality reduction techniques. It has been applied to many pattern recognition researches such as face recognition (Abbas, Safi, & Rijab, 2017). In addition to PCA, LDA has also been widely employed to reduce the features size (Moon, Choi, Kim, & Pan, 2015). LDA enhances the class separability by providing the most discriminant projection of features from a high dimensionality to a low dimensionality, however the performance of LDA drops down when the samples number is smaller than the dimensionality (Fukunaga, 2013).

To address the problem of feature selection optimization, evolutionary algorithms have been widely applied to decrease the features size and enhance the selection process. The most used algorithms are GA and PSO (particle swarm optimization).

In Zhao, Fu, Ji, Tang, and Zhou (2011) GA is used to select the features and set the parameters for SVM (support vector machine) which will significantly influence the classification accuracy. The proposed has been evaluated on several real-world datasets, according to the results, the fusion of GA and SVM has improved and the accuracy and the processing time.

In Guo, White, Wang, Li, and Wang (2011) they propose GA feature selection (GAFES) in software product line (SPL). GA is used to minimize or maximize an objective function such as a total cost. According to the results, GAFES proves that it can achieve better results than other feature selection methods and in less time.

In Huang (2012) they propose a stock selection method using GA and support vector regression (SVR). GA is used to optimize the model parameters and select the optimal set of input variables to the SVR model. The proposed GA-SVR has significantly outperformed the benchmark.

In Gheyas and Smith (2010) They proposed SAGA for feature selection which is a hybridization between simulated annealing (SA) and GA. SA is good at avoiding local minimum and GA has a high convergence rate. According to the results, the hybrid method outperforms other algorithms on real and synthetic datasets.

In Kabir, Shahjahan, and Murase (2011) they proposed a hybrid GA and local search. They embedded a local search technique in GA to fine tune the search. The local search aims to distinguish the nature of features based on their correlation information. They evaluate the performance of their approach on 11 real-world classification datasets according to the results their method performs better than other well-known feature selection algorithms.

In Suguna and Thanushkodi (2010) they combine GA with KNN. Instead of choosing the k-neighbors among all the samples, GA is used to select the k neighbors immediately and classify the new samples by computing the distance.

In Sidorov, Brester, Minker, and Semenkin (2014) they use self-adaptive multi-objective GA for feature selection and probabilistic neural network for classification. The results show that the proposed approach increased the emotion recognition accuracy by up to 26.08%.

In Ghamisi and Benediktsson (2015) they proposed an integration of a GA and PSO, where the accuracy of SVM is used as fitness function. According to the results performed on indian pines hyperspectral dataset, their method can select relevant features and achieve high accuracy within a reasonable time.

In Oreski and Oreski (2014) a hybrid GA and neural networks (HGA-NN) is proposed as feature selection for credit risk assessment. They use a real-world credit dataset to evaluate the performance of HGA-NN classifier.

In Anbarasi, Anupriya, and Iyengar (2010) a hybrid of GA-SVM is used to select a relevant subset of bands mass in order to classify hyperspectral images, which can be applied to land cover investigation and target detection.

Datasets in medical domain consist of much more disease measurements than records. Therefore feature selection is necessary to reduce redundancy and the memory space used to stock the disease measurements.

In Li, Wu, Wan, and Zhu (2011) in order to reduce the number of tests that the patient needs to take, GA is employed to determine which tests contribute more toward the diagnosis of heart diseases, which will reduce the number of tests automatically. Classification by clustering, decision tree and naive Bayes is then applied to predict the diagnosis. In Inbarani, Azar, and Jothi (2014) a hybridization of PSO is proposed, PSO based quick reduct (PSO-QR) and PSO based relative reduct (PSO-RR) are used to optimize the selection of features for diseases diagnosis.

In addition to GA and PSO which are widely employed for feature selection, ant colony optimization (ACO) also has been employed to reduce the size of features and select an optimal subset of features. In Kashef and Nezamabadi-pour (2015) Advanced Binary ACO (ABACO) is proposed for feature selection. The features form a fully connected graph where each node consists of two sub-nodes, one to select and the other to deselect the feature. At each tour, each ant has to visit all the nodes on the graph. At the end of the tour, every ant will have a vector of zeros and ones where 1 means select and 0 means deselect the corresponding feature. In Chen, Chen, and Chen (2013) they propose an alternative way to be traversed by the ants, for n features instead of using a fully connected graph they use a directed graph with only O(2n) arcs instead of $O(n^2)$ arcs.

Framework

In this section, we present the proposed framework which consists of three basic steps, the first one is feature extraction, in which, each image is transformed into a feature vector which is supposed to be a representative of the original face image, the next step is feature selection which is the main aim of our study, in this step, the size of feature vectors will be significantly reduced by eliminating the irrelevant features and keep only a small subset of features that best represents the face emotion. These vectors are then used in the classification step to segregate the features into different facial emotion classes. Fig. 1 shows the framework of our approach.

Before extracting the feature vectors, the images go through a preprocessing step, in which the region of interest is extracted from the original images. The absolute face area is extracted from the background. Extracting the face area is necessary to get more accurate emotion detection. Then the images are resized to 100∗100 pixels and normalized to zero mean. Viola-Jones algorithm is employed to detect and extract the face area (Viola & Jones, 2004). The adaboost algorithm used in viola-jones face detection aims to construct a strong classifier by laniary combining weak classifiers. Viola-jones is a robust and fast face detection technique and it can be applied for real-time applications. After extracting the face area we applied canny edge (Canny, 1987) detection operator.