Informatics in Medicine Unlocked 9 (2017) 58–67

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

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

# Multimodal detection of concealed information using Genetic-SVM classifier with strict validation structure

Ehsan Darestani Farahani, Mohammad Hassan Moradi *

*Biomedical Engineering Faculty, Amirkabir University of Technology, Tehran, Iran*

ARTICLE INFO

*Keywords:*
Genetic-SVM
Strict validation structure
Concealed information test (CIT)
Event-related potential (ERP)
Autonomic response

ABSTRACT

The main idea of multimodal detection of concealed information is based on a comparison between the physiological responses of a subject to the probe and irrelevant items. The purpose of the present study was to extend the previous study by using a new machine learning structure to discriminate guilty and innocent subjects. The new machine learning structure was designed based on utilizing genetic support vector machine (GSVM) as a classifier. Evaluation of GSVM was performed through a strict validation structure to prevent from optimistic result. But this method entailed the problem of finding several optimal parameters in one optimization problem. To solve this problem, an innovative approach based on "the most frequent result" was introduced, and therefore the general optimal parameters in each feature subset were obtained.

The classification accuracy of GSVM based on ERP feature was approximately 90%. The best accuracy among autonomic features and their combinations was around 95.45% and among combinations of ERP and autonomic features was around 93%. In comparison with previously used methods on the same data, GSVM had higher accuracy in all of the feature subsets.

The new presented machine learning structure made experimental conditions closer to real conditions by using strict validation and introducing general optimal parameter. On the other hand, the higher accuracy of GSVM shows the better SVM's ability to find complex patterns and solve classification problems as compared with the other methods. Moreover, the crucial role of Genetic algorithms to find the optimal parameters shouldn't be ignored.

## 1. Introduction

Detection of deception has great importance from legal, moral and clinical aspects and therefore has been considered by scientist from different research area. Currently, the polygraphic tests are the most widely used technique for the quantitative discrimination between deceptive and truthful responses [2,33]. The common physiological measures of polygraph systems, which reflect the function of the autonomic nervous system (ANS), are respiration activity, heart rate (HR) and electrodermal response [35]. These physiological measures are limited and indirect view of complex underlying brain processes in a deception procedure. To overcome this problem some alternative approaches have been studied. Investigation of brain function in deception procedure is one of these approaches. Some of neurophysiological signals such as, Functional Magnetic Resonance Imaging (fMRI) [34,36] and event related potentials (ERP) have recently been investigated for the possible application of them in detection of deception [1,17,39]. The ERP based studies reported more satisfying results compared to other neurophysiological studies [39]. P300 wave is the most important component of ERP, which is used in many studies. Previous ERP-based studies reported that the amplitude of P300 is a reliable index of deception detection [17,38,40]. In the recent years, the simultaneous measurement of ERP and autonomic responses (i.e. classic physiological measures of polygraph systems) was exploited to improve the result of the detection of deception.

The concealed information test (CIT) is an effective psychophysiological test for detection of concealed information about crime [9]. In this test two types of items are severally presented to the suspect. One of these items which related to the feature of the under investigation crime, is called probe item and the other one is called irrelevant. In CIT, the differences between the physiological responses of suspect to probe and irrelevant items are surveyed. A guilty subject who deceive his/her knowledge about crime shows different physiological responses to probe and irrelevant items, whereas an innocent subject without detailed

* Corresponding author.
*E-mail address:* mhmoradi@aut.ac.ir (M.H. Moradi).

http://dx.doi.org/10.1016/j.imu.2017.05.004
Received 5 February 2017; Received in revised form 23 April 2017; Accepted 9 May 2017
Available online 12 May 2017

information about the crime shows similar physiological responses to both items [27].

ERP measurement that mirrors cognitive processing and autonomic responses that reflect a function of ANS can complement each other and yield higher accuracy in the result of CIT. However, combining ERP and autonomic measurements in one experiment entails some difficulties such as habituation problem and overlapping in skin conductance responses (SCR) [5,16]. The studies exploiting simultaneous measurement can be divided into studies which used short inter stimulus interval (ISI) and large number of stimulus presentations [5,31] and studies with long ISI and low number of stimulus presentations [20,29]. In the second category most studies did not find a significant difference in P300 amplitude between probe and irrelevant. They suggested that exploiting long ISI is a possible reason of these non significant ERP responses and using short ISI could be a solution [20]. In the first category, the studies achieved incremental validity from the combination of measurements, while the result of discrimination based on autonomic measurements were less than other autonomic-based studies.

Classification of subjects into innocent and guilty group is one of the important practical aspects of a deception detection system. So far, the classification methods which have been used in multimodal lie detection studies were too limited. In the most of the studies logistic regression has been used for data classification and stepwise logistic regression to test for the incremental validity [5,20,31]. Previously, we implemented a multimodal CIT test with linear discriminant analysis (LDA) and logistic regression as a classifier [16]. It seems that other classification methods such as support vector machine (SVM) can be helpful in this case. Strong robustness to noise, generalization to the unseen data and the unique solution are the main advantages of SVM [6]. So far, SVM has been used in many research areas and demonstrated a good performance as a pattern recognition method [32]. Because SVM is a parametric classifier, finding optimal parameters of this classifier has a crucial role in its performance. In some studies heuristic methods such as genetic algorithm were used to find the optimal parameters for SVM which led to introduce the Genetic-SVM classifier [7,23].

In the classification problem in addition to training data set, a testing data set is needed for testing the generalization performance of classifier. In the simplest case the original dataset is divided into training set and testing (holdout) set. If the amount of data is limited the data dividing will be very important and there will be a lot of variation in the performance estimate for different partitions of the data to form training and test sets. Using by cross-validation (CV) can be ensuring that every sample from the original dataset has the same chance of appearing in the training and testing set. In the basic approach, called k-fold CV, the data set is split into k smaller sets and train k times, treating a different set as the holdout set each time. At last the generalization performance is obtained by averaging over the values computed in each iteration [13]. So the estimated performance is less sensitive to the partitioning of the data.

When some heuristic methods are used to find optimal parameters of a classifier or optimal feature subset, the performance of classifier is used as a cost function to guide the search. This performance can't be considered as generalized performance because in this way the parameter can be tweaked until the classifier performs optimally on test set and make overfitting. A few studies have reported the CV estimates that were used to guide the search as a final estimate of performance, thus achieving overly optimistic results [25]. To solve this problem a strict validation by means of independent data for calibration and validation is needed. It means when the optimal parameter have found based on performance of classifier on calibration data, the performance on separate validation data should be reported as generalized performance. In the case of limited amount of data which a CV is used to guide the search, another holdout sets generated by an outer loop of cross-validation can be used to estimate of generalized performance [22,25,37]. By using a CV in outer loop, an optimal parameter is obtained for each iteration. The obtained optimal parameters aren't necessarily equal. So a question arises that which one is the best one as a general solution. In this paper an innovative approach based on the most frequent answer was proposed to find an optimal parameter among the set of optimal parameters.

In our previous study short ISI and large number of stimulus presentation were chosen for getting acceptable results from ERP measure. To overcome the limitations of this paradigm on autonomic measures, some solutions were suggested and investigated [16]. The purpose of the present study was to extend the previous study by using a new machine learning structure. The new machine learning structure was designed based on GSVM as a classifier with strict validation technique and an innovative approach to introduce the general optimal parameters (GOP).

The paper is organized as follows. At first the procedure of our experiments and the characteristics of the collected data will be explained. The extracted feature will be described afterwards. In the next step, the basic concept of SVM and the structure of proposed GSVM will be explained. Finally, the performance of this novel method will be compared with the previous classifier.

## 2. Materials and methods

### 2.1. Participants

Fifty-two healthy subjects (40 m, 12f; mean age 22.5 ± 3.5; all right-handed; all had normal or corrected vision) voluntarily participated in this study. All participants were gifted a small gold coin (its price was around US $12) after the experiment.

### 2.2. Design and procedure

Participants were asked to randomly select one envelope out of two, containing the scenario of excrement. Experimenter mentioned that the envelopes are containing the instruction of a guilty or an innocent scenario. However, the experimenter had selected one scenario beforehand and put the same instruction in both envelopes.

After selection of the scenario, the participants had to perform scenario based on the instruction. The guilty scenario was based on stealing a gold coin (its price was around US $12, and it was hidden in a wallet) and a cellphone from a personal locker. The innocent scenario consisted of washing dirty cups put in the kitchen sink. Finally, participants were asked to go to the lobby and wait for experimenter. After 7 min, the experimenter went to lobby and informed participant about a crime and encouraged them to win a gold coin. The participants try to get the gold coin while they do not know that the two envelopes were the same. It is expected that these arrangements prevent the participants from inattentiveness.

In current study, a CIT test with five blocks was used (Fig. 1). In each block, 7 pictures related to the crime sense (e.g. coin, cellphone, …) were severally presented. Each picture performed the role of one type of stimuli (i.e., target, probe, 4 irrelevants, and a null event) and was presented 7 times (except for null picture 6 times). The pictures were presented based on a pseudo random sequence called M-sequence with varying ISI, 2.3–2.7 s (for more detail please see Farahani and Moradi [16]).

The participants were asked to attend to all presented pictures and press right click of the mouse when they recognize the target picture and left click for all other pictures.

26 participants were randomly selected to perform the guilty scenario. To improve the sensitivity of CIT test, the innocent subjects did not have prior knowledge about the guilty scenario [45]. A few experiment results were excluded due to misdoing of the protocol or noisy measurements. Finally, 23 guilty subjects and 21 innocent subjects were used for subsequent analyses.

### 2.3. Physiological recording

Physiological recordings were made in a silent environment (laboratory), without the presence of other people.