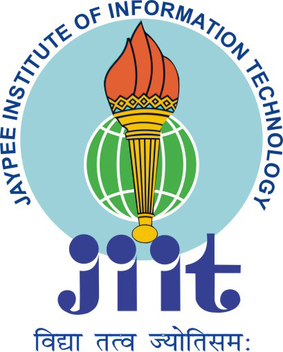
TERM PAPER

*Mid Evaluation Report* - *Major Project*



**Deception Detection of familiar/unfamiliar faces using EEG signals**

         04.10.2019

Ajay Kumar   16103177

Akash Gupta 16103201

Saurabh Gupta 16103065

Supervisor :  Mrs. Indu Chawla

External Panel :  Mrs. Anuja Arora , Mr. Prashant Kaushik

# LIST OF PAPERS:

1. Better than random? A closer look on BCI results
2. Classification of EEG Signals by using Support Vector Machines
3. Analysis of EEG Signals for Deception Detection

# 4. Evaluation of P300 based Lie Detection Algorithm

# 5. Lie Detection Based EEG-P300 Signal Classified by ANFIS Method

6. Wavelet analysis for EEG feature extraction in deception detection

7. Deception Detection of EEG-P300 Component Classified by SVM Method

8. An Improved Approach for EEG Signal Classification using Autoencoder

9.Classification of EEG Signals by using Support Vector Machines

# PAPER - 1

**Better than random? A closer look on BCI results**

By : Gernot R. Müller-Putz, Reinhold Scherer, Clemens Brunner, Robert Leeb, Gert Pfurtscheller

Published In: International Journal of Bioelectromagnetism

**Summary:**

Brain-computer interface (BCI) is a collaboration between a brain and a device that enables signals from the brain to direct some external activity, such as control of a [cursor](https://whatis.techtarget.com/definition/cursor) or a prosthetic limb.

This paper is a basic comparison between Theoretical approach and Simulation(Practical) approach of BCI for the case of 2-class paradigm.

The theoretical approach states that the probability of correctly classified trial in a 2-class paradigm of N trails follows a binomial distribution with p = 0.5, i.e, both classes are equally likely to occur. In this context a BCI experiment consisting of trials (50 per class), the expected chance level would be at exactly 50 correctly classified trials (with equally probable classes). If the reported accuracy of a classifier is 59 correctly classified trials (or alternatively, 59%), it is straightforward to see that this probability does not lie within the theoretical limits of 40.39% and 59.61% (for a confidence of ). Thus, it can be assumed that the given classifier does not significantly differ from a random one.

The theoretical results for BCI of a 2-,3-,4- and 8-class are 50%, 33.3% 25% and 12.5% respectively. However, the Simulation results vary from the theoretical for N = 80, the results of 2-,3-,4-,8-class BCI are 57.5%, 39.6%, 29.7% 15.2% for a significance level α=5%.

Concluding, we want to take into account the proposed considerations and to check their results also in relation to the real level of chance and not only to the theoretical one.

# PAPER - 2

**Classification of EEG Signals by using Support Vector Machines**

By : K. Sercan Bayram, M. Ayyüce Kızrak, Bülent Bolat

Published In: ResearchGate

**Summary:**

In this work, EEG signals were classified by support vector machines to detect whether a subject’s planning to perform a task or not. Various different kernels were utilized to find the best kernel function and after that, a feature selection process was realized.

EEG recordings of four different mental states was classified by using five classifiers. In this work, the best result was obtained by using a resilient back propagation method as 95%. With the help of multilayer perceptron(MLP) it was recognised whether a person was sleeping or was awake by using EEG signals.

The five different metal activities were also classified with Support Vector Machine with maximum accuracy of 72%. The accuracy for using a quantum neural networks was 81.33%. Another classifier used was neural networks had an accuracy of 80%. The ID3 classifier had an accuracy of 71%. The final classifier used was fuzz rough ID which correctly classified 76.2% of the data.

In the first step of the work, the dataset classified by linear and nonlinear SVMs, and the best kernel function was determined. Based on the results, the best choice is radial basis function (RBF) kernel for σ=1.2 with 71.43% accuracy. By applying the SFS on data, the result had the accuracy of 72.53%. The SBE gave the best result as 74.73%. At last t-scores and p-values were calculated. The accuracy range of t-score varied from 67 to 71.43% and that of p-value varied from 70.33 to 71.43%.

# PAPER - 3

**Analysis of EEG Signals for Deception Detection**

By : Roshani J. Khandelwal, Juilee D. Mahajan, Ujjwala P. Bombatkar, Snehal G. Badhe

Published In: International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

**Summary:**

Prominent deception detection approaches include the standard polygraph which monitors the signature changes in autonomic responses and the cognitively more central EEG. .Deception detection is the practice of attempting to determine whether someone is lying. The common signs of deceptive behavior are body language, emotional gestures and contradiction, interactions and reactions, verbal context and content,facial micro-expressions, change of topic etc. These deception techniques are used by police, forensic psychologists, security experts and other investigators to help prevent them from being a victim of fraud or scams and other deceptions.

Throughout history, it has often been assumed that lying is accompanied by a change in the body’s physiological activity. The polygraph is a set of equipment that accurately measures various sorts of bodily activity such as heart rate, blood pressure, respiration, and palm sweating.

Lying causes a conflict between lie and the truth within the brain. The increased activity can be detected by F- MRI which records brain activity by identifying changes in brain blood flow and the metabolic rate. The radar based procedure which could perform remote, unobtrusive, non-invasive and stealthy lie detection is when an UWB radar pulse passes through the human thorax it gets echoed back by the cardiac structure i.e. the heart wall. Heart rate variability is the physiological phenomenon of variation in the time interval between heartbeats i.e. the variation in the beat-to-beat interval. HRV is also an indicator of the emotional arousal.

Several morphological features were extracted to know the various parameters and distinguish

truth from lie telling. Classification is done by Euclidean distance method, which will calculate the minimum value between the vectors to display the output. The output is generally displayed with a message as “lie EEG signal” or “true EEG signal”.

The conceptual study indicates that it is feasible to identify the basic lie detection.EEG signals will be used for the lie detection because this is non invasive, cheap, and are the direct results of the electric activity inside the brain.

# PAPER - 4

# Evaluation of P300 based Lie Detection Algorithm

By : Syed Kamran Haider, Malik Imran Daud, Aimin Jiang, Zubair Khan

Published In: IEEE

**Summary:**

Lie detection is used to verify people’s value. Earlier methods were unable to accurately determine mental behaviour.A researchable analysis gives us an idea that, with the help of EEG signals we can easily monitor the psychological variations, brain activities and deception detection related features But lately, we can get an idea of psychological behaviour using Electroencephalogram(EEG). We use an EEG to detect brain signals.

We extract the desired features from the brain signals acquired through sixteen electrodes using various extraction techniques. Then, they have implemented linear discriminant analysis (LDA) classification technique to differentiate the positive and negative samples from the signals obtained from sensors in order to achieve a decision for either guilty subject or innocent subject accordingly.

Twenty subjects (15 males and 5 females, ages between 20-25 years) that were generally universities students and all had good health with stable psychological behaviour participated in the study.

 They recorded data of over 15 to 20 subjects. In one scenario, some precious items (e.g., jewelry, cash, smart phones, etc.) were placed in front of some subjects. Those subjects were not informed about the scenario and the subject can be any random person. Another person called as Subject-2 intended to perform this test come and stole that item that was placed in front of person called as Subject-1. It can be any person but suggestively that can be his/her close friend. As this test was performed in university, university authorities would handle the case of that stolen item. This was done only to put our subject under pressure for lie and truth response.

First some obvious questions were asked that everybody would answer the obvious answers, i.e., the true answer. The brain signals were recorded with the sample rate of 128 samples per second. The recorded signals which were digitized and amplified can be monitored offline with the help of MATLAB software. Pass band filter were used with the range of 0.3-30 Hz prior to data analysis. For P300-based study about Guilty Knowledge Test (GKT) the previous mentioned pass band filter range were used [9]. The EEG signals from EPOC headset are itself digitized because headset contains the built-in analog to digital converter (ADC).

After applying pass band filter on EEG signals, each continuous subject record is divided into single sweep as per the times known by stimulus presentation [10]. The total length of each single sweep is 1000ms, and contains 128 samples. Then the signal pattern recognition technique includes special features extraction and features selection. After that apply classification method on the signals to assess the detection rate. It should be noticed that, in all cases related to P300 ERP the Pz is the prime location where P300 can be monitored maximal and therefore visual related experiments were performed only on Pz data.

 The EEG signal is in time domain and the energy is distributed. The optimal features are suppressed with the noise. In order to dig out the features, EEG signal is observed under the signal energy in the form of time domain or frequency domain. The frequency domain analysis is the best for those features utilized in the mental task related to EEG signals.

LDA is a dimensionality reduction technique which is commonly used for the supervised classification problems. It is used for modeling differences in groups i.e. separating two or more classes. It is used to project the features in higher dimension space into a lower dimension space.

For example, we have two classes and we need to separate them efficiently. Classes can have multiple features. Using only a single feature to classify them may result in some overlapping as shown in the below figure. So, we will keep on increasing the number of features for proper classification. Achieved 85% accuracy in detecting lie and truth subjects.

# 

# PAPER - 5

# Lie Detection Based EEG-P300 Signal Classified by ANFIS Method

By : Arjon Turnip, M. Faizal Amri, M. Agung Suhendra, and Dwi Esti Kusumandari

Published In: Springer

**Summary:**

The lie detector is an instrument that is often discussed or researched by scientists and experts. Because of a number of problems posed by lies and frauds, which be able to lead to criminal activities, lie detector needs to be improved. Many subjects or suspects of criminal activities to lie when questioned by authorities. This shows the importance of tools that can differentiate between a subject who is lie or not.

Once the signals recording was complete, the continuous EEG data from each subject were inspected and filtered for artifacts using band-pass filter and Independent Component Analysis (ICA), respectively. Parts of the signals that contained noises by task-irrelevant movement or artifact be cut by band-pass filtered using 0.1 Hz and 30 Hz cut-offs [17] and then the noises were removed by ICA.

In this research, discrete wavelet transform (DWT) is used as an extraction method. The reason why the wavelet transform has been selected because the component of ERP signal-to-noise ratio (SNR) is low and not stationary

The DWT uses multi filter banks and special wavelet filters for the analysis and reconstruction of signals. The DWT provides a compact representation of a signal in time and frequency that can be computed efficiently. The method calculates the wavelet coefficients at discrete intervals of time and scale instead of at all scales.

An Adaptive Network Fuzzy Interference System (ANFIS) is used as a classifier after signals extraction.The results of signal from three stimuli responses which are produced through signal processing, response from P stimuli has the most important information in determining whether subjects are lying or not. Before we got the signal features that affected by P stimuli, preprocessing signals and feature extraction had been through.

The ANFIS method applied at the features classification step has the advantage of much less training time is achieved. The results indicated that the existing method in this article had great result for lie detection. The ANFIS method is able to separate lying subjects from honest subjects based on EEG-P300 signals with an accuracy of 64.27%.

# PAPER - 6

# Wavelet analysis for EEG feature extraction in deception detection

By : Anna Caterina Merzagora, Scott Bunce, Meltem Izzetoglu and Banu Onaral

Published In: 28th IEEE

EMBS Annual International Conference

New York City, USA

**Summary:**

# By interfacing the brains signals directly, it is possible to design brain computer interfaces to control devices without mechanical interfaces. Especially direct controlling the prosthetic organs is very important for disabled people. The most common brain activity monitoring device is electroencephalogram (EEG).

A primary emphasis in this study was to examine the capacity for physiological measures to differentiate among the cognitive elements of truth and deception, i.e. the knowledge that one is lying. The task was designed to elicit high motivation to escape detection, but to minimize participants’ anxiety about being deceptive. To accomplish this end, the task was framed as a form of poker-like card game in which it is socially acceptable to “bluff,” or to lie, minimizing feelings of anxiety about lying.

 Participants were given a total of 5 cards, four of which (one from each suit) were face-up on the computer screen (the ‘hand’). Participants were informed that the identities of these four face-up cards, as in some forms of poker, were known by the participants, as well as the researchers. Participants were then asked to choose a fifth card from among three sealed envelopes, each of which contained a playing card which they kept in their hand (‘target’ card) and $50. Participants were informed that only they knew the identity of this card, and the experimenter would be attempting to learn the identity of this card by alternately presenting a series of cards, asking the question

“Do you have this card?”, and examining their brain responses. They were told that if they were successful in concealing the identity of the card, that would be able to keep the $50, in addition to their participation remuneration ($25).

 Quadratic B-spline wavelets were used in the wavelet analysis due to their near optimal time frequency localization properties. Moreover, their waveform is similar to the waveforms to be detected in the EEG signal; hence extraction of EEG components is more likely to be successful.

 The aim of this preliminary study was to investigate the ability of wavelet domain features obtained from the EEG differentiate truth from deception during a low anxiety task. The results revealed that wavelet coefficients corresponding to beta waves were found to differentiate when subjects were telling the truth versus when they were lying.

PAPER - 7

**Deception Detection of EEG-P300 Component Classified by SVM Method**

By : K. Sercan Bayram, M. Ayyüce Kızrak

Published In: IEEE

**Summary:**

By interfacing the brains signals directly, it is possible to design brain computer interfaces to control devices without mechanical interfaces. Especially direct controlling the prosthetic organs is very important for disabled people. The most common brain activity monitoring device is electroencephalogram (EEG). The ability to measure noninvasively the related brain activity of lying within an individual subject could offer a significant improvement over currently available tools to detect deception

A vast variety of approaches to the classification of quantitative features from an EEG signals. SVM have become extremely successful discriminative approaches to pattern classification and regression problems. In other words, SVM is a technique used to obtain the most probable hyperplane to separate two classes. It is done by measuring the hyperplane’s margin and determines its maximum point. Margin is defined as distance between the corresponding hyperplane and the nearest pattern from each class.

To enhance signal noise ratio of P300 components, the independent component analysis (ICA) method was adopted to separate non-P300 (i.e. artifacts) and P300 components from

every single trial. Then the P300 waveforms with high SNR were reconstructed. And then group of features based on time, frequency, and amplitude were extracted from the reconstructed

P300 waveforms. Finally, two different class of feature samples were used to train a support vector machine classifier because it has higher performance compared with several other classifiers.

PAPER - 8

An Improved Approach for EEG Signal Classification using Autoencoder

By : Abhijith V Nair*∗*, Kodidasu Murali Kumar*y*, Jimson Mathew

Published In: IEEE

Publishing Date: January 30, 2018

**Summary:**

Electroencephalography (EEG), being one of the major brain signal using in Brain-Computer Interface (BCI) applications and in many others also. Comparing to other brain signals, it is very easy to acquire and very cheap. Also, there will be no risk for the user while acquiring the EEG signals and also it is a non-invasive technique. EEG signals are used for clinical as well as research purposes.

Analyzing the Event Related Potentials (ERPs) triggered while the subject is exposed to different faces [4], N170 and P2 waves gave the degree of familiarity effects with small amplitude signals after 170ms and 250ms respectively.

Independent Component Analysis (ICA) as the signal processing tool. Due to the low Signal-to-Noise ratio (SNR) value of EEG components, using of ICA enhance the conditioning thus by

reducing the noise component. By finding the suitable linear combinations of mixed variables helps for the calculation of independent components. ICA is not a commonly used one,

comparing to Principal Component Analysis (PCA). But it is better for signals having multiple recordings at the same time.

Based on the property of non-linear feature extraction technique of autoencoders, we designed a

classifier for classifying the familiarity and unfamiliarity of the images.

PAPER - 9

**Classification of EEG Signals by using Support Vector Machines**

By : K. Sercan Bayram, M. Ayyüce Kızrak

Published In: IEEE

Publishing Date: 2013

**Summary:**

By interfacing the brains signals directly, it is possible to design brain computer interfaces to control devices without mechanical interfaces. Especially direct controlling the prosthetic organs is very important for disabled people. The most common brain activity monitoring device is electroencephalogram (EEG).

Support vector machine (SVM) is a statistical learning theory based classification method. For a given two-class linearly separable classification problem, SVM tries to find a hyperplane which separates the input space with a maximum margin. The first group which called as filters deals each features independently form theothers. Due to their nature, filters are fast and computationally cheap. On the other hand, filter methods can’t interpret therelations between features, and this disability limits the filters’ performance. Wrappers consist of a searching algorithm and a classifier. The search algorithm searches new solutions through the feature space while the classifier produces a fitness function to the search algorithm. The most common

algorithms are genetic algorithms, particle swarm optimization,

The algorithm starts with the entire dataset. In every step, existing features are removed from the dataset one by one. The feature that gives the least decrease on the accuracy is

excluded from the dataset. The algorithm continues until the stopping criterion is reached To raise the accuracy and to find the most relevant features, four different feature selection algorithms applied to the dataset. To find the most suitable kernel, different SVMs were

trained with original data. The data divided into two equal parts as training and test data.

All of the experiments realized ten times, and averaged to find the accuracies. In the first step of the work, the dataset classified by linear and nonlinear SVMs, and the best kernel function was determined.