**Design of a Bio-signal Based Stress Detection System Using Machine Learning Techniques**

**ABSTRACT**

This work illustrates the creation of a machine learning-based stress detection system using certain readily available bio-signals in the human body. The typical definition of stress is the disruption of psychological balance. One of the main fields of biomedical engineering study is stress detection, which is important because accurate stress detection may easily stop many psychological and physical issues, such as arrhythmia or irregular cardiac rhythm. There are many bio-signals that can be used to determine stress levels, such as the ECG, EMG, respiration, GSR, and others. These signals exhibit distinctive variations when under stress. Because of the readily available recording (i.e., the market currently offers numerous mobile clinical grade recorders) and ECG feature extraction methodologies, the ECG was chosen as the primary option in this study. Another benefit of ECG is the ability to detect respiratory signal information from ECG, known as EDR (ECG derived Respiration), without the need for a separate sensor device for measuring respiration.

ECG signals have recognizable characteristics, and collecting the data is inexpensive. For the creation of the model, we deduced the RR interval, QT interval, and EDR features from the ECG. The "drivedb" database from Physionet served as the training dataset and validation for the construction of a supervised machine learning (SVM) approach in MATLAB. SVM was chosen for classification since the labelled data can be divided into two categories: "stressed" and "non-stressed." By altering the feature quantity and Kernel type, several SVM model types were validated. Our results showed an accuracy level of 98.6% with Gaussian Kernel function and using all available features (RR, QT and EDR), which also emphasizes the importance of respiratory information in stress detection through Machine Learning.

**Keywords:** Stress Detection, Arrhythmia, ECG Derived Respiration (EDR), Machine Learning, MATLAB.

**LITERATURE SURVEY**

**[1] Fevre, Mark Le; Kolt, Gregory S.; Matheny, Jonathan, "Eustress, distress and their interpretation in primary and secondary occupational stress management interventions: which way first?" Journal of Managerial Psychology, 2006, 21 (6): 547 -565.doi:10.1108/02683940610684391**

Purpose to develop an argument for the retention of secondary approaches to stress management (those that focus on the individual within the organization) as first interventions, prior to the employment of primary approaches (those that focus on the organization's processes and structures). This is based on a reconsideration of eustress versus distress and a review of current empirical evidence on the effectiveness of stress management interventions. Design/methodology/approach Major empirical studies and reviews are critically reviewed and placed within a theoretical framework derived from both early and more recent work in the field. Findings There is little empirical evidence on which to base recommendations for organization‐based stress management interventions as first or sole approaches and therefore the value of these as first or sole approaches is questioned. Instead secondary, individual‐focused, approaches are recommended as first‐line interventions prior to the adoption of organization‐based interventions. Practical implications. In practice secondary stress management approaches are currently most common. Broader primary approaches appear to have excellent theoretical support and a growing body of supportive literature and accompanying recommendations for employment. We suggest, however, that secondary approaches be employed prior to the introduction of primary methodologies within a client organization. Originality/value. This paper provides a review and framework for interpreting/understanding the research on the effectiveness of stress management interventions and makes recommendations relevant to practitioners in the field.

**Summary:** Studied about the Eustress, distress and their interpretation in primary and secondary occupational stress management interventions.

**[2] Brodsky MA, Sato DA, Iseri LT, Wolff LJ, Allen BJ. “Ventricular Tachyarrhythmia Associated with Psychological Stress.” JAMA. 1987; 257:2064-7:**

Psychological stress has been reported to be a risk factor for sudden cardiac death in individuals both with and without underlying structural heart disease. From a group of 80 patients presenting with life-threatening ventricular tachyarrhythmia, six were identified without underlying structural heart disease. Five of these six patients experienced marked psychological stress. Each of these five patients underwent arrhythmia evaluation, demonstrating recurrent rapid monomorphic ventricular tachycardia related to changes in tone of the sympathetic nervous system. Subsequently, solitary beta-adrenergic blocker therapy was given to each patient. During therapy, four of the five patients had a marked reduction of both arrhythmia and symptoms during a follow-up ranging from 29 to 49 (mean, 38) months.

**Summary:** Studied about the Ventricular Tachyarrhythmia Associated with Psychological Stress.

**[3] Brunckhorst CB, Holzmeister J, Scharf C, Binggeli C, Duru F. “Stress, depression and cardiac arrhythmias.” Ther Umsch. 2003; 60:673-81:**

Stress can exert adverse effect on cardiovascular health. Psychosocial stress adversely affects the autonomic homeostasis. This in turn can result in metabolic abnormalities, inflammation and dysfunction of endothelium. Changes in the autonomic homeostasis can be a major trigger for ventricular tachyarrhythmias. Increased sympathetic nervous activity can cause increased proarrhythmic repolarization instability leading to spontaneous ventricular arrhythmias. During stress-induced autonomic nervous system activity, the heart rate rises and the heart rate variability indices like low frequency power falls before the onset of ventricular tachycardia. Psychological stress has been shown to induce T wave alternans, which in turn predicts future ventricular tachyarrhythmia events. Fluctuations in T wave amplitude after psychological stress are predictive of subsequent arrhythmic events.

**Summary:** Studied about Stress, depression and cardiac arrhythmias.

**[4]** **Leenhardt A, Lucet V, Denjoy I, Grau F, Ngoc DD, Coumel P. “Catecholaminergic polymorphic ventricular tachycardia in children. A 7-year follow-up of 21 patients.” Circulation. 1995; 91:1512-9:**

We observed 21 children (mean +/- SD age, 9.9 +/- 4 years) at the time of the diagnosis who had no structural heart disease and a normal QT interval on routine ECG. They were referred for stress- or emotion-induced syncope related to ventricular polymorphic tachyarrhythmias. The arrhythmia, consisting of isolated polymorphic ventricular extrasystoles followed by salvoes of bidirectional and polymorphic tachycardia susceptible to degeneration into ventricular fibrillation, was reproducibly induced by any form of increasing adrenergic stimulation. There was a familial history of syncope or sudden death in 30% of our patients. On receiving therapy with the appropriate beta-blocker, the patients' symptoms and polymorphic tachyarrhythmias disappeared. During a mean follow-up period of 7 years, three syncopal events and two sudden deaths occurred, probably due to treatment interruption.

**Summary:** Studied about Catecholaminergic polymorphic ventricular tachycardia in children a 7-year follow-up of 21 patients.

**[5] Taggart, Peter & Boyett, Mark & Logantha, Sunil Jit & Lambiase, Pier. (2011). “Anger, Emotion, and Arrhythmias: From Brain to Heart. Frontiers in physiology.” 2. 67. 10.3389/fphys.2011.00067:**

Strong emotion and mental stress are now recognized as playing a significant role in severe and fatal ventricular arrhythmias. The mechanisms, although incompletely understood, include central processing at the cortical and brain stem level, the autonomic nerves and the electrophysiology of the myocardium. Each of these is usually studied separately by investigators from different disciplines. However, many are regulatory processes which incorporate interactive feedforward and feedback mechanisms. In this review we consider the whole as an integrated interactive brain–heart system.

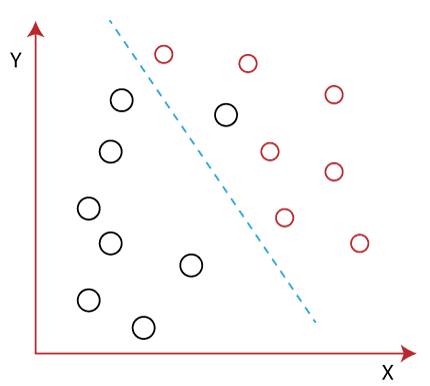
**SUMMARY:** Studied about Anger, Emotion, and Arrhythmias: From Brain to Heart. Frontiers in physiology.

**EXISTING METHOD**

This systematic review groups and summarizes the algorithms used to detect the state of stress and the physiological and behavioral features used to feed these algorithms, associated with each type of stressor. Method: We conducted a systematic literature review following the PRISMA-statement in seven databases. The studies considered for this review were those that used physiological or behavioral response to detect stress state. Results: 27 publications (29 independent studies) were included in the review. Stress detection accuracy ranged from 54% using a Decision Tree (DT) to 100% using Linear Discriminant Analysis (LDA). 72.4% studies used psychological stressors and 27.6% physical stressors to generate the stress state. Conclusions: The behavioral response has not been widely studied in stress detection and may be the key to identify which stressor is generating a particular stress state.

**Linear Discriminant Analysis (LDA)**

Linear Discriminant Analysis (LDA) is one of the commonly used dimensionality reduction techniques in machine learning to solve more than two-class classification problems. It is also known as Normal Discriminant Analysis (NDA) or Discriminant Function Analysis (DFA).



**Fig: Linear Discriminant Analysis**

This can be used to project the features of higher dimensional space into lower-dimensional space in order to reduce resources and dimensional costs. In this topic, "Linear Discriminant Analysis (LDA) in machine learning”, we will discuss the LDA algorithm for classification predictive modeling problems, limitation of logistic regression, representation of linear Discriminant analysis model, how to make a prediction using LDA, how to prepare data for LDA, extensions to LDA and much more. So, let's start with a quick introduction to Linear Discriminant Analysis (LDA) in machine learning.

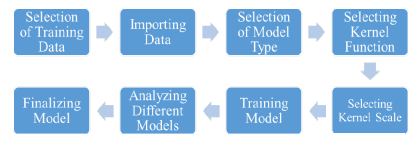
**Disadvantages:**

* Computationally complex.
* Less features generation for training.

**PROPOSED METHOD**

In this study, stress has been identified using supervised machine learning. The ECG characteristics QT interval, RR interval, and ECG Derived Respiration were chosen to identify stress. The model will determine whether the subject is stressed or calm by examining them. We have selected Support Vector Machine (SVM) for the classification and detection of stress because there are exactly two class labels in our data (stressed and not stressed or relaxed).

Both continuous and discrete data sets can be utilized with the Support Vector Machine (SVM). However, Classification Technique is mostly utilized for discrete data sets. This approach involves plotting the data in an n-dimensional plane, where n is the total number of features. After that, categorization is carried out by identifying the hyperplane that two classes.



**Fig: Flow of Proposed Method**

Figure shows the Data and decision variables processed for Model Training. First, the model was trained with three SVM types (Linear, Quadratic and Cubic) using the default kernel functions and the cross validation scheme was chosen as the holdout validation with a degree of 50% in the Classification Learner App. All of the three features i.e. QT interval, RR interval, and ECG Derived Respiration (EDR) were used. By analyzing the scatter plots, confusion matrix, and ROC curves the accurate model can be chosen. Model accuracy depending on different types of SVM is represented in Table. Further modifications can be done with the model to find accurate results.

To analyze the effect of different features, the model was then trained using only one feature to understand if it is possible to detect stress accurately. A significant drop in the model accuracy was found if we use only one feature as shown in Table. If we train the model with only two features by unchecking any of the three features from “Feature Selection” option of Classification Learner and train the model again, the model accuracy becomes low compared to the previous one. Model accuracy differences by training with only two features can be noticed from Table.

Combination of QT interval and RR interval was found to give an accuracy of almost 95%. Other models failed to generate comparable amount of accuracy using combination of two features (Table). Almost all of the model-based stress detection techniques until now have used only one feature, mostly RR interval to detect stress as found in the literature. But their performance is not acceptable for biomedical applications (i.e. Accuracy level less than 95%). Therefore, in this study we trained the model with QT interval, RR interval and ECG Derived Respiration (EDR) separately and analyzed the results. We have chosen EDR (ECG Derived Respiration) as an alternate for respiration signal as they show similar properties and be used as successfully in many previous studies.

Kernel Function” can also be tuned to improve the model performance. Therefore, we have trained the models with different kernel functions. Linear and Quadratic kernel functions did not show good results. Only Gaussian and Cubic kernel function types showed promising results in detecting stress using all three features of ECG as shown.

**ADVANTAGES AND APPLICATIONS**

**Advantages:**

* The more features we use, the more accurate the model becomes.
* The extraction of more features is key in reducing training time.
* SVM classifier takes less time for a better training.

**Applications:**

* Machine Learning
* Deep Learning
* Digital Signal Processing
* Medical Signal Processing

**Software & Hardware Requirements:**

**Software:** Matlab 2020a or above

**Hardware:**

**Operating Systems:**

* Windows 10
* Windows 7 Service Pack 1
* Windows Server 2019
* Windows Server 2016

**Processors:**

Minimum: Any Intel or AMD x86-64 processor

Recommended: Any Intel or AMD x86-64 processor with four logical cores and AVX2 instruction set support

**Disk:**

Minimum: 2.9 GB of HDD space for MATLAB only, 5-8 GB for a typical installation

Recommended: An SSD is recommended A full installation of all MathWorks products may take up to 29 GB of disk space

**RAM:**

Minimum: 4 GB

Recommended: 8 GB

**Learning outcomes:**

* Introduction to Matlab
* What is EISPACK & LINPACK
* How to start with MATLAB
* About Matlab language
* Matlab coding skills
* About tools & libraries
* Application Program Interface in Matlab
* About Matlab desktop
* How to use Matlab editor to create M-Files
* Features of Matlab
* Basics on Matlab
* What is an Image/pixel?
* About image formats
* Introduction to Image Processing
* How digital image is formed
* Importing the image via image acquisition tools
* Analyzing and manipulation of image.
* Phases of image processing:
* Acquisition
* Image enhancement
* Image restoration
* Color image processing
* Image compression
* Morphological processing
* Segmentation etc.,
* How to extend our work to another real time applications
* Project development Skills
  + Problem analyzing skills
  + Problem solving skills
  + Creativity and imaginary skills
  + Programming skills
  + Deployment
  + Testing skills
  + Debugging skills
  + Project presentation skills
  + Thesis writing skills