Abstract— Stress is a major problem in our world today motivating objective understanding of how average individuals respond to stress in a typical activities. The main aim for this paper is to determine whether stress can be recognized using individual-independent computational models from sensor based stress response signals induced by films with typical stressful content. Another aim is to determine whether a consumer electroencephalogram (EEG) sensor device, which is portable, less obtrusive and relatively inexpensive, can be used for stress recognition. A support vector machine and an artificial neural network based models were developed to recognize stress using various physiological and physical signals. The models produced stress classification with 95% accuracy. Using the data obtained from the consumer device, the models produced stress classification with 91% accuracy. Statistical analysis of the results showed that the classification results from the physiological and physical signals are not statistically different to the results from the consumer device implying that the consumer device can be used for recognizing stress in typical virtual environments.

Keywords—stress recognition; support vector machines; artificial neural networks; genetic algorithms; physiological signals; physical signals; EEG; films

I. INTRODUCTION

Stress is part of everyday life and it is widely accepted that stress which leads to less favorable states (such as anxiety, fear or anger) is a growing concern for people and society. The term, stress, was coined by Hans Selye. He defined it as “the non-specific response of the body to any demand for change” [1]. Stress is the body’s reaction or response to the imbalance caused between demands and resources available to a person. It is seen as a natural alarm, resistance and exhaustion [2] system for the body to prepare for a fight or flight response to protect the body from threats and changes. When experienced for longer periods without being managed, stress has been widely recognized as a major growing concern because it has the potential to cause chronic illnesses (e.g. cardiovascular diseases, diabetes and some forms of cancer) and increase economic costs in societies, especially in developed countries [3, 4]. Benefits of stress research range from improving day-to-day activities, through increasing work productivity to benefitting the wider society - motivating interest, making it a beneficial area of research and posing some difficult technical challenges for Computer Science.

Various computational methods have been used to objectively define and classify stress to differentiate conditions causing stress from other conditions. The methods developed have used simplistic models formed from techniques like Bayesian networks [5, 6], decision trees [7] fuzzy models [8] and support vector machines [9]. These models have been built from a relatively smaller set of stress features than the sets used in the models in this paper.

Stressful events or emergency situations cause dynamic changes in the human body, which can be observed by changes in the body’s response signals. These response signals are involuntarily caused by the Autonomic Nervous System (ANS). The ANS is made up of the Sympathetic Nervous System (SNS) and the Parasympathetic Nervous System (PNS). When the body is under stress, activity in the SNS increases and dominates the activities produced by the PNS, which changes the body’s response signals. The response signals obtained from non-invasive methods that reflect reactions of individuals and their bodies to stressful situations have been used to interpret degrees of stress. These measures have provided a basis for defining stress objectively.

Stress response signals used in this paper fall into two categories – physiological and physical signals. Physiological signals that have been used for stress analysis include galvanic skin response (GSR) [10, 11], electrocardiogram (ECG) [12] and blood pressure (BP) [13]. As defined in [14], physical signals are changes made by the human body that can be generally seen by humans without the need for equipment and tools that need to be attached to individuals to detect general fluctuations. However, sophisticated equipment and sensors using vision technologies are still needed to obtain physical signals at sampling rates sufficient for data analysis and modeling like the ones used in this paper. Physical signals include eye gaze and pupil dilation signals.

GSR is a measurement of the flow of electricity through the skin of a person. When the person is under stress, the activity in the SNS causes an increase in the moisture on the skin, which increases the flow of electricity. As a result, it increases skin conductance [5]. Conversely, the skin conductance is reduced when the individual becomes less stressed. The fluctuations in skin conductance are recorded as changes in GSR.

ECG signals show electrical activity produced by an impulse of ions flowing through cardiac muscles, which dissipates into the region around the heart with diminished
amounts spreading around the surface of the body. The main electrical signals are produced by cardiac cells depolarising and repolarising. Depolarising occurs due to the flow of ions accompanying atrial heart muscle constriction which results in a P wave. The impulse then travels through the ventricles of the heart causing septal depolarisation, early ventricular depolarisation followed by late ventricular depolarisation. This series forms a QRS wave, the dominant wave. After the completion of depolarisation, ventricular cells repolarise by restoring it to resting polarity, resulting in a T wave. An example of a stress feature in ECG signals is the decrease in R-R intervals [12].

BP signals show the pressure exerted on the walls of blood vessels due to blood circulation and varies with a systolic (maximum) and a diastolic (minimum) pressure. An increase BP has been related with increase in stress [15].

GSR, ECG, BP, eye gaze and pupil dilation signals have been used in some smaller combinations to detect stress in literature [5, 16, 17] but this proposed combination for stress analysis in virtual environments has not been reported so far. We use this combination of sensor signals in this paper and refer to them as primary stress signals.

Hundreds of stress features can be derived from primary stress signals to classify stress. However, this set of features can include redundant and irrelevant features which may outweigh the more effective features showing stress patterns. As a consequence, this could cause a classifier to learn weaker stress patterns and produce lower quality classifications. Since this paper is dealing with sensor data, some features may suffer from corruption as well. In order to achieve a good classification model that is robust to such potential features that may reduce the performance of classifications, appropriate feature selection methods must be developed and adopted by the classifiers. A genetic algorithm (GA) could be used to select subsets of features for optimizing stress classifications. A GA is a global search algorithm and has been commonly used to solve optimization problems [18]. The search algorithm is based on the concept of natural evolution. It evolves a population of candidate solutions using crossover, mutation and selection methods in search for a population of a better quality. GAs have been successfully used to select features derived from physiological signals [19, 20].

This paper also investigates the potential for electroencephalogram (EEG) signals sourced using the Emotiv EPOC consumer device [21] as a primary signal for stress. EEG captures neural activity in the brain. It records complex electrical waveforms at the scalp formed by action electrical potentials during synaptic excitations and inhibitions of dendrites. Research shows that relationships exist between brain activity and stress [22-26]. EEG signals are generally captured by equipment that is obtrusive and requires a relatively higher degree of intrusion (e.g. EEG sensors with conductivity gel needs to be attached to particular positions of a subject’s head), relatively long set up time and experimenters that are experienced with the use of the equipment. These factors contribute to lower number of participants in experiments recording EEG signals in Human Centered Computing experiments and motivated us to use the Emotiv EPOC consumer device. Emotiv is portable, nonintrusive, easier to use and cheaper – the Emotiv headset with EEG sensors (and costs approximately $300 USD). Unlike conventional EEG recording equipment, which has multiple wires that needs to be managed, Emotiv is wireless. However, Emotiv has fewer sensors (14 channels) and a lower sampling rate (128 Hz) than the conventional EEG recording equipment which can have from 32 to 256 sensors and a sampling rate of at least 240 Hz to 16,000 Hz [27, 28]. The channels for Emotiv are AF3, AF4, F3, F4, F7, F8, FC5, FC6, P3, P4, P7, P8, T7, T8, O1, and O2 based on the International 10-20 standard. Previous work has used signals from Emotiv for analysis of affective states [29]. Our work includes a comparison of the performances for stress classification for models developed from Emotiv signals with performances for models developed from primary stress signals sourced by sensitive sensors. This will help to determine whether future stress research could use signals sourced from the less obtrusive Emotiv sensors for computationally modeling stress.

This paper presents computational models of stress for observers watching stressful films and investigates whether models formed from signals sourced using a consumer EEG device have the capacity to recognize stress. It describes the method that was used to collect stress data and develop computational models for recognizing stress patterns captured in response signals observed from individuals while they watched stressed and non-stressed film clips. It details an experiment conducted to collect stress response signals using multiple sensors and EEG signals from the consumer EEG device from experiment participants who watched films with stressful and non-stressful content and human-reported data. Several approaches for stress recognition of film watchers are developed, compared and discussed including a method for selecting features from hundreds of features derived from the response signals. Performances for models formed from the primary signals of stress for stress recognition are compared with the models formed from the EEG signals. The paper concludes with a summary of the findings and suggests directions for future work.

II. DATA COLLECTION FROM THE FILM EXPERIMENT

Twenty five undergraduate students were recruited as experiment participants. The participant cohort was made up of 15 males and 10 females between the ages of 18 and 30 years.

Before the start of each experiment, the lab room was set up with the temperature set at 22°C. Acquisition and Inquisition computers with required applications started, consent forms and assessment sheets placed at appropriate locations in the room, and all sensor equipment initialized and ready to be attached to a participant. The lab room is made up of two smaller rooms – Acquisition and Inquisition rooms. The experiment instructor controlled the data acquisition, films that participants watched and the room settings of the Inquisition room from the Acquisition room. The participant watched the film clips and had their measurements recorded in the Inquisition room.
Each participant had to understand the requirements of the experiment from a written set of experiment instructions and what was involved in the experiment with the guidance of the experiment instructor. After providing their consent to participate in the experiment, the experiment instructor attached ECG, GSR, BP and EEG sensors to the participant and calibrated the eye tracking sensors. The instructor started the films, which triggered a blank screen with a countdown of the numbers “3”, “2” and “1” transitioning in and out slowly with one after the other. The reason for the countdown display and the blank screen was for participants to move away from their thoughts at the time and get ready to pay attention to the films that were about to start. This approach was used in experiments for similar work done in literature [30]. Subsequent to the countdown display, a blank screen was shown for 15 seconds, which was followed by a sequence of film clips with 5 second blank screens in between. The film clips were made up of stressed and non-stressed films and each film was approximately one minute in length. After the films finished, the participant was asked to do an assessment on the films to validate the film labels. In total, the experiment took approximately thirty minutes. An outline of the process of the experiment for a participant is shown in Figure 1. After the experiment, the experiment instructor helped the participant to remove the sensors that were attached to the participant.

Film clips were categorized by the type of environment they created. Stressed films had stressful content in the direction towards distress, fear and tension whereas the non-stressed films had content that created an illusion of meditation or soothing environments. There were three stressed films and three non-stressed films. For consistency, the films were displayed on a 1050 x 1680 pixel LCD Dell monitor and positioned at the same location of the computer screen with a surround sound system for each participant.

Results from the experiment survey validated the film classes. In addition, the film classes were also validated by another forty people giving a total of 65 people who did the

Figure 1. The process that experiment participants followed during the film experiment

Figure 2. The Inquisition room for the film experiment where experiment participants did the experiment
Analysis of survey responses is a common method used in literature to validate stress classes for tasks [31]. Participants found the films that were labeled stressed created stressful environments and films labeled non-stressed created environments that were not stressful with a statistical significance of p < 0.001 according to the T-test and the distribution of the survey question responses for stress.

Each participant had physiological and physical signal measurements taken over the experiment time. The physiological and physical sensor signals (which we refer to as primary stress signals) captured during the experiment were GSR, ECG, BP, eye gaze and pupil diameter signals. Biopac ECG100C, Biopac GSR100C and Finapres Finger Cuff systems were used to take ECG, GSR and blood pressure recordings at a sampling rate of 1000 Hz. The Biopac GSR100C system sourced GSR data through disposable EL507 GSR electrodes, which were placed on the first (or index) and fourth (ring) fingers of the participant’s left hand. The sensors had to be attached to the participant 5 minutes before data acquisition so that appropriate signals could be recorded. For acquiring ECG signals, the Biopac ECG100C system was connected with disposable EL503 ECG electrodes, which were placed on the participant’s left and right wrists. Further, the Finapres Finger Cuff system was attached to the participant’s third (or middle) finger of the left hand to acquire BP signals. Eye gaze and pupil dilation signals were obtained using Seeing Machines FaceLAB system with a pair of infrared cameras at 60 Hz. Additionally, EEG signals from the Emotiv EPOC device was obtained as well at a sampling rate of 128 Hz.

A schematic diagram of the equipment setup for the film experiment is provided in Figure 3. The experiment instructor controlled the film clips that the experiment participant watched from the acquisition computer. The computer was also used to control the participant’s response signal acquisition and synchronize the signals while the films were shown to the participant. Acquisition software applications for eye tracking and EEG signals required relatively greater amounts of computer memory and processing. In order to ensure that it did not affect the acquisition of the other signals, there was a computer designated for eye tracking signal acquisition and another computer for EEG signals.

There were other signals that were derived from the primary stress signals to form other stress response signals. These signals included the heart rate variability (HRV) signal, which was calculated from consecutive ECG peaks and another popular signal used for stress detection [12, 32]. Data from various frequency bands were extracted from EEG and HRV signals, which have been used in literature for stress analysis. There are four main frequency band categories used to analyze EEG signals and they are presented in TABLE I. The band categories are Beta, Alpha, Theta and Delta. Each band category represents some state for a person. Beta and alpha waves represent conscious states of a person whereas theta and
delta waves signify unconscious states. Rapid beta wave frequencies (and due to the decrease in alpha wave frequencies) have been found to indicate stress [24-26].

Generally, low frequency (LF) and high frequency (HF) bands of HRV have been used to analyze stress in literature [8]. LF and HF are known to reflect SNS and PNS activities respectively. LF was defined in the frequency range between 0.04 Hz and 0.15 Hz [31] and HF was defined in the range between 0.15 Hz and 0.5 Hz [13].

Features were derived from the primary stress signals and EEG signals. Statistics (e.g. mean, standard deviation, skewness, kurtosis and measures listed in [33]) were calculated for the signal measurements for each 5 second interval during the experiment time. Measures such as the number of peaks for periodic signals, the distance an eye covered, the number of forward and backward tracking fixations, and the proportion of the time the eye fixated on different regions of the computer screen over 5 second intervals were also obtained. The EEG feature set also included measurements for Hjorth parameters and fractal dimensions, which are generally used in analyzing EEG signals.

Hjorth parameters are time-based characteristics of an EEG signal and the three Hjorth parameters are the Activity, Mobility and Complexity parameters. Suppose \( x \) is an EEG signal with values for \( N \) equally spaced timestamps. Then the Activity parameter is the variance for an EEG signal and is defined by

\[
Activity(x) = \frac{\sum_{n=1}^{N} (x_n - \bar{x})^2}{N}
\]  

(1)

The Mobility parameter is a measure of the signal mean frequency. Given that \( x' \) is the derivate for \( x \), then the Mobility parameter is defined by

\[
Mobility(x) = \frac{Activity(x')}{\sqrt{Activity(x)}}
\]  

(2)

The Complexity parameter is a measure of the deviation of the EEG signal from the shape of the sine signal and is defined by

\[
Complexity(x) = \frac{Mobility(x')}{Mobility(x)}
\]  

(3)

Fractal dimension measures of an EEG signal provide information of the space filling and self-similarity and can be approximated using the following

\[
Fractal Dimension (x) = 1 + \frac{\log (L)}{\log (2(N - 1))}
\]  

(4)

where

\[
L = \sum_{n=2}^{N} \left( \frac{(x_n^* - x_{n-1}^*)^2 + (n - (N - 1))}{N} \right)
\]  

(5)

and

\[
x^* = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}
\]  

(6)

The statistic and measure values formed the stress feature set. There were 215 features derived from the primary stress signals and 1119 features were derived from the EEG signals sourced by the 14 Emotiv sensors. Features from the primary stress signals and EEG signals were used as inputs to the stress recognition models.

III. STRESS CLASSIFICATION MODELS

The stress classification models were based on a support vector machine (SVM) and an artificial neural network (ANN). A GA and SVM hybrid (GA+SVM) and a GA and ANN hybrid (GA+ANN) models were also developed with the aim to use features, which were more relevant for stress recognition and reduce feature redundancy to optimize stress classification.

A. Support Vector Machine Based Classification Models

SVMs have been widely used in literature for classification problems including classifications based on physiological data [34, 35]. Provided a set of training samples, a SVM transforms the data samples using a nonlinear mapping to a higher dimension with the aim to determine a hyperplane that partitions data by class or labels. A hyperplane is chosen based on support vectors, which are training data samples that define maximum margins from the support vectors to the hyperplane to form the best decision boundary. This contributes to the resistance to data overfitting and helps to generalize classifications well.

Despite the useful characteristics, SVMs are still not robust to feature sets with redundant and irrelevant features in classification. As a consequence, hybrids of SVM with GA were used to reduce the redundant and irrelevant features and to determine whether the hybrid improved the quality of the classification.

The SVM based stress classification models developed were:

<table>
<thead>
<tr>
<th>Band Category</th>
<th>Frequency Range</th>
<th>Person’s State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alpha</td>
<td>Relaxation</td>
</tr>
<tr>
<td>Beta</td>
<td>13 – 30</td>
<td>Alertness or anxiety</td>
</tr>
<tr>
<td>Alpha</td>
<td>8 – 13</td>
<td>Dream sleep or phase between consciousness and drowsiness</td>
</tr>
<tr>
<td>Theta</td>
<td>4 – 8</td>
<td>Dream sleep or phase between consciousness and drowsiness</td>
</tr>
<tr>
<td>Delta</td>
<td>0.5 – 4</td>
<td>Coma or deep sleep</td>
</tr>
</tbody>
</table>

TABLE I. EEG FREQUENCY BAND CATEGORIES
TABLE II. IMPLEMENTATION SETTINGS FOR THE GENETIC ALGORITHM

<table>
<thead>
<tr>
<th>GA Parameter</th>
<th>Value/Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>population size</td>
<td>100</td>
</tr>
<tr>
<td>number of generations</td>
<td>2000</td>
</tr>
<tr>
<td>crossover rate</td>
<td>0.80</td>
</tr>
<tr>
<td>mutation rate</td>
<td>1/(length of the chromosome)</td>
</tr>
<tr>
<td>crossover type</td>
<td>MATLAB’s Scattered Crossover</td>
</tr>
<tr>
<td>mutation type</td>
<td>MATLAB’s Uniform Mutation</td>
</tr>
<tr>
<td>selection type</td>
<td>MATLAB’s Stochastic Uniform Selection</td>
</tr>
</tbody>
</table>

- SVM: a SVM classification model that was provided all the features in the stress feature set as input to recognize stress patterns
- GA+SVM: a SVM that was provided features selected by a GA as input to recognize stress patterns

B. Artificial Neural Networks Based Classification Models

ANNs, inspired by biological neural networks, have capabilities for learning patterns to recognize characteristics in input tuples by classes. An ANN is made up of interconnected processors, known as artificial neurons, which are connected by weighted links that pass signals between neurons to learn the relationship between input and output tuples. In this paper, feed-forward ANNs trained using backpropagation were used. The layers and neurons in each layer define the topology of a feed-forward ANN. Two topologies were used, which differed only on the number of inputs. The ANN based stress classification models were:

- ANN: an ANN classification model that used all the features in the stress feature set as input to recognize stress patterns just like SVM
- GA+ANN: an ANN that used inputs selected by a GA to recognize stress patterns

The MATLAB adapt function was used for training the ANNs on an incremental basis. Each network was trained using the Levenberg-Marquardt algorithm for 1000 epochs or until the magnitude of the gradient for the mean square error was less than $10^{-6}$. The mean square error is the average squared error determined from the actual output of the network and the expected output. The ANNs had 7 hidden neurons and one neuron in the output layer.

C. A Genetic Algorithm for Optimization of Stress Recognition

A GA was used to optimize the stress features and determine whether the SVM and ANN classification model systems were able to better capture stress patterns. It was used to investigate whether feature selection improved the quality of stress classifications. The GA search evolved a population of subsets of features using crossover, mutation and selection methods in search for a population of subset of features that produced a better quality stress classification. A subset of features is referred to as an individual or chromosome. The quality for each chromosome in the population was defined by the quality of classifications produced when a classifier was provided with the features encoded by the chromosome.

The initial population for the GAs was set up to have all the features. The number of features in the chromosomes varied but the chromosome length was fixed. The length of a chromosome was equal to the number of features in the feature space. A chromosome was a binary string where the index for a bit represented a feature and the bit value indicated whether the feature was used in the classification.

The GAs in the hybrid classification systems were implemented with the settings provided in TABLE II.

IV. RESULTS AND DISCUSSION

The computational classification models were tested on the stress data obtained from the film experiment. Features derived from the primary stress signals were provided to the SVM, ANN, GA+SVM and GA+ANN classification systems. The feature values were normalized by participant to reduce the effect of individual bias. Performances of the classification models were calculated using 10-fold cross validation. The classification results are provided in TABLE IV.

Stress patterns in the film data set were best recognized with the GA hybrids according to the stress recognition measures. When the classification model systems were provided all the features derived from the primary stress signals, the accuracy and the F-score were at least 0.24 lower than systems with the GA feature selection. Classification results show that the performances of SVM and ANN were similar. Similarly, the performances of GA+SVM and GA+ANN in stress classification were similar.

The GA in the GA hybrid classification systems selected ECG features for measurements of the LF and HF frequency bands, gradients, rate of change, mean and variance for GSR, ECG and BP, rate of change of the distance covered by eye gazes calculated by eye fixations, and the rate of change of the diameter size of the eye pupil. This feature subset had stress data on how the experiment participants reacted to the different films and provided improved capability for the classification systems to capture better stress patterns in the film experiment data set and recognize stress better.

Further, EEG features were provided as input to the SVM, ANN, GA+SVM and GA+ANN classification systems as well. The classification results based on 10-fold cross validation are shown in TABLE III.
Like the performance of the classification systems on the primary stress signals, GA hybrids produced better stress recognition rates and F-score values than the systems that used all features to develop a model. This shows that the GA hybrids captured stronger stress patterns in their models. The stress recognition measures for the systems without the GA was at least 0.16 lower than the GA hybrid systems. In addition, SVM and ANN produced similar stress recognition rates and the GA hybrids produced classifications with similar stress recognition rates.

The GAs selected measures for EEG fractal dimensions, and variance, mean and skewness features of the EEG signals and signals in the Beta and Alpha frequency bands for stress classification. The features selected by the GAs had stronger relationship with stress and enabled the classifiers in the classification system to better capture stress patterns in the film experiment EEG data set.

Statistical analyses were conducted on the results and in accordance with the Student’s T-test the GA hybrid systems produced significantly better stress recognition rates than the systems without the GA with $p < 0.01$. Analysis of the stress recognition rates produced by systems on the feature sets derived from primary stress signals and the EEG signals showed no statistical difference for the two feature sets ($p > 0.1$). This suggests that the Emotiv system has the capability to capture stress signals and has the potential to be used for stress recognition in environments where intrusion needs to be minimized for easier stress data acquisition.

V. CONCLUSION AND FUTURE WORK

Computational models of stress for observers of films were developed based on SVMs and ANNs using a real-world stress data set formed from a film experiment. The data set was made up of physiological and physical signals as well as EEG signals sourced by a consumer device – the Emotiv system. Stress patterns in the data were successfully captured by the SVM and ANN models. The hybrids of GA with SVM and GA with ANN performed significantly better than the stress recognition systems without a GA with higher stress recognition results. Analysis of the results produced by the stress recognition systems on the data sourced by the consumer EEG device shows potential for the device as a tool for stress analysis for the type of work presented in this paper. Some extension to this work will include optimal sensor selection for the consumer EEG device and modeling of online stress analysis. Future work will also include stress analysis for observers of other types of environments such as more natural environments with real-life settings with more suitable sensors that are non-obtrusive to suit day-to-day data capture and stress analysis.

REFERENCES


