

Assessment of Driver Stress from Physiological Signals collected under Real-Time Semi-Urban Driving Scenarios

Rajiv Ranjan Singh*

*Department of Electrical and Electronics Engineering, Birla Institute of Technology & Science (BITS), Pilani
Rajasthan, India, 333031*

**Corresponding Author: rrsingh@pilani.bits-pilani.ac.in; rrsingh70@gmail.com*

Sailesh Conjeti

*School of Medical Science and Technology, Indian Institute of Technology Kharagpur
West Bengal, India, 721302*

Rahul Banerjee

*Department of Computer Science, Birla Institute of Technology & Science (BITS), Pilani
Rajasthan, India, 333031*

Received 22 April 2013

Accepted 5 September 2013

Abstract

Designing a wearable driver assist system requires extraction of relevant features from physiological signals like galvanic skin response and photoplethysmogram collected from automotive drivers during real-time driving. In the discussed case, four stress-classes were identified using cascade forward neural network (CASFNN) which performed consistently with minimal intra- and inter-subject variability. Task-induced stress-trends were tracked using Triggs' Tracking Variable-based regression model with CASFNN configuration. The proposed framework will enable proactive initiation of rescue and relaxation procedures during accidents and emergencies.

Keywords: Wearable Driver Assist Systems, Physiological Signals, Affective State, Stress-Trends, Neural Networks.

1. Introduction

Road accidents are one of the major causes of deaths around the world. Over 1.2 million people die each year on the world's roads, and between 20 and 50 million suffer non-fatal injuries, amounting to an estimated loss of US\$ 518 billion due to road traffic injuries and costs governments between 1%-3% of their gross national product (GNP) globally¹. Over 90% of the world's fatalities occur on the roads of low-income and middle-income countries¹. Among the 37,261 fatalities and 2.35 million injuries in USA in 2008, 64% were drivers, 27% passengers and remaining others (4% motorcyclists, 3% pedestrians and 2% pedalcyclists)². In USA, over 95% single vehicle run-off-road (ROR) crashes between 2005 and 2007, critical reasons were attributed to drivers including driver performance errors (27.7%), followed by driver decision errors (25.4%), critical non-performance errors (22.5%) and recognition errors (19.8%)³. Crashes attributed to vehicles were only 1% and due to environment only 1.1%. The findings also showed that driver's inattention, fatigue and hurriedness were the most influential factors³.

In real-life traffic situations, driving becomes stressful due to the frequent occurrence of events and incidents, thereby affecting the concentration of drivers⁴. These events are sequential maneuvers (like stopping for a light, changing

lanes, putting on the brakes), whereas the incidents are frequent but unpredictable (like near misses, frustration due to overtaking or not getting a pass etc.). These events and incidents are sources of physiological responses attributed due to extreme physiological reactions, emotional reactions and irrational thoughts leading to stress⁴. Unacceptable levels of stress, fatigue and on-road distractions deteriorate driver's performance⁵ and may lead to temporal loss of concentration, risk assessment capability and vehicular control, often inviting road accidents.

Lisetti and Nasoz⁵ in their study showed that there exists a strong relationship between the galvanic skin response (GSR) signal and frustration whereas heart rate is more related to anger and fear. The GSR increased for a frustrated person whereas the heart rate increased for a feared person while it decreased for an angry participant. Katsis et al. used support vector machines (SVMs) and adaptive neuro-fuzzy inference system (ANFIS) to classify emotions as high stress, low stress, disappointment and euphoria and attained 79.3% and 76.7% classification rates for SVM and ANFIS respectively⁶. Healey and Picard⁷ collected biosignal data to monitor the driver's stress during real on-road experiments. They recommended that future vehicles should have provision to integrate vehicular infrastructure along with body-worn sensor configuration for on-road driver stress monitoring.

To address this issue, our work is a part of a research initiative which aims to enhance driver safety by designing a custom wearable computing fabric as a component of a context-aware driver assistant system which can save loss of precious lives by the way of providing fast yet credible real-time alerts to the drivers and their coupled cars⁸. It is proposed that in such systems if the decision making modules were trained on the basis of naturalized data collected from real-time driving scenarios, machines will be able to adapt and detect the stressful and alarmable situations effectively.

Therefore we decided to include *affective state* (also called *emotional state* or the *sentic state* to assess the mental and physical stress¹⁰) and *stress-trend*¹¹ (discrete events and incidents observed during drive⁴) detection approaches for the stress level assessment of drivers. The salient features of our stress level detection approach includes: (a) first work to the best of our knowledge focusing a developing country like India in terms of scenario design, subject population and road settings, (b) physiological data collected in real-time driving scenarios modeled the stress contributing factors into a multiclass problem instead of a binary class⁹, (c) an exhaustive set of physiological features (39 statistical, syntactic and spectral) were extracted representing driver's current physiological state, (d) instead of less number of subjects and a one-fold classification^[6-7], we used more subjects and performed analysis on single as well multi-turn drive data, and (e) six neural network 4-class classifiers were evaluated instead of generalizing a single classifier.

2. Data Collection Methodologies

In the present work, we collected data from 20 drivers, majorly belonging to Pilani in Shekhawati region of Rajasthan, India using body-mounted physiological sensors during multiple fixed route segmented on-road driving experiments.

2.1. Participants

The chosen drivers form a representative population ranging from novice to professional, with an average age (at the time of data collection) of 33.65 ± 9.93 years (Min. 21 and Max. 58) and with a mean experience of 11.60 ± 8.39 years (Min. 3 and Max. 37) Among them, 15% were casual, 55% (11 drivers) with short distance and remaining 30% drivers with long distance driving experience. Predominantly male drivers were considered because of non-availability of female professional drivers in Pilani who volunteered for the study and also due to the low representation of women in the Indian professional vehicle driver workforce (less than 3%) and exposure to roads¹². It was hence decided that it is better to adopt a custom design approach targeting the men driver

population. Design factors to be adopted for female drivers will be incorporated in the future experiments.

2.2. Physiological Signals, Sensors and Experimental Setup

Physiological features have been extracted from signals like electrocardiogram (ECG / EKG), electromyogram (EMG), galvanic skin response (GSR), respiration and electrocardiogram (EKG / ECG) and classified for detecting stress level or emotions of automotive drivers^[5,6,7]. The mostly used physiological parameters have been (a) heart rate (HR) and heart rate variability (HRV), both derived from ECG and (b) skin conductance responses (SCR) and other related features which account for sudden stress responses, derived from GSR. The photoplethysmogram (PPG) signal has been established as an alternate for ECG to derive several features including HR, HRV, percentage oxygen saturation (SpO2) and Respiration Rate¹³.

We employed the Mind media B V's Nexus-10 device¹⁴ alongwith Biotrace+ software for data acquisition. Initially, four sensors viz. a II lead ECG sensor, a body-worn clip-on Nonin Pulse Oximeter for PPG and SpO2 signals, an abdominal respiration belt and GSR Velcro electrodes, as shown in Figure 1(a), attached to one of the four channels of the device to collect the physiological data. During data collection in driving scenarios in several of the driving sessions it was noticed that ECG and respiration signals were corrupted beyond recovery which may have happened due to the dynamic chest movement. Therefore we concentrated on PPG and GSR signals for the analysis. PPG signal is a permanent signal¹⁵, which can be used to extract multiple parameters¹³ like HR, SpO2, Respiration Rate etc. GSR signal is an induced signal, as certain derived feature responses directly correlates due to triggering of stressful events¹⁵. During stimulus morphological changes are observed and the signal has a tendency to gradually deviate the baseline, which is suitable for identifying stress¹⁶. The PPG signals were collected at a sampling frequency of 128 Hz and GSR at 32 Hz. PPG was further down sampled to 32 Hz for inter-signal compatibility during the offline processing stage.

2.3. Data Collection Scenarios

The total data acquisition time-line is segmented into five scenarios comprising of two relaxed scenarios and three driving scenarios as shown in Table 1. In the two relaxed scenarios, Pre-driving (Pr-dr) and Post-driving (Po-dr), data was collected in laboratory setup while drivers were relaxed, depicted in Figure 1 (b). In the three driving scenarios the data was collected, in university campus named Relaxed-driving (Rx-dr), on city roads and highway named

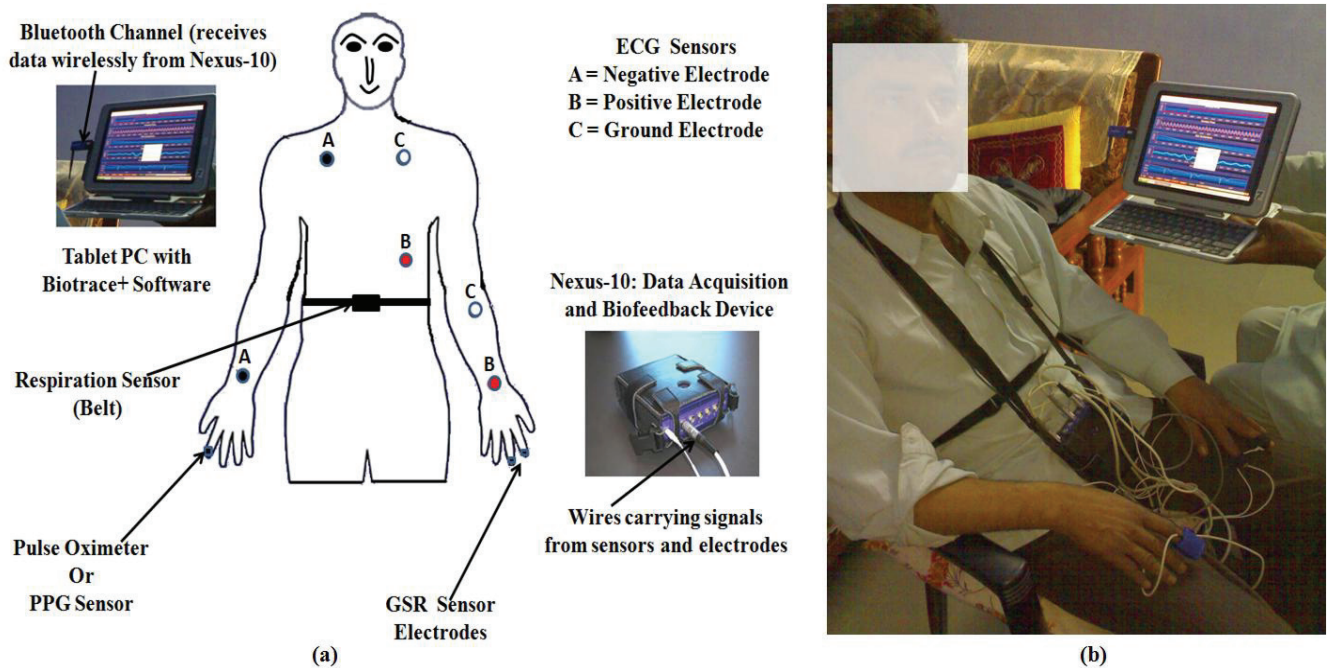


Figure 1. Experimental Setup: (a) Typical Sensor Placement Locations and Data Acquisition (b) Pre-driving Data Collection Scenario.

Busy-driving (By-dr), and in university campus named Post-driving (Po-dr). The total drive time for each driver lasted for nearly 24 minutes covering a distance of approximately 11.5 kilometers. The advantages of on-road uncontrolled data collection in different driving scenarios over simulated environments may include (a) training classifier on real-time data makes it robust to noise, motion artefacts, device errors etc. (b) effect of factors like environmental, vehicle's characteristics and driver's physiological conditions will be considered (c) correlation of stressful events will be more accurate than in simulated conditions¹¹. During preliminary analysis presented in our previous work it has been successfully established using multivariate analysis of variance (MANOVA) that the choice of data collection scenarios are significant and raw signals like GSR, Respiration and %SpO2 were also found to be statistically important¹⁷.

Drivers have to maneuver through several right-turns, left-turns, sharp turns (left, right and circular) immediate stops, busy market areas with high vehicle and pedestrian density, handling bad stretches of streets, negotiating ill-designed or non-marked speed breakers, abrupt lane change by a neighboring vehicle, jaywalkers etc. These maneuvers and stressful events lead to repeated distractions, thereby increasing the cognitive workload on drivers leading to stress and fatigue⁴. In this paper such on-road events would henceforth be referred as *stress-trends* and their automatic detection would enable the wearable computer to activate and respond in accident prone spells of driving. Therefore, stress-trend time marker based annotation was carried on simultaneously by a secondary experimenter during real-time data collection. The data collected under these scenarios was converted into a compatible data format for necessary signal processing procedures in an offline workstation using MATLAB[®].

Table 1. Data Collection Scenarios

Scenarios	Location	Route Length	Time (min.)	Speed-Limit (Kmph)	Stress-Trends	Traffic Density		
						Ped.	2-Wh./Bi.	4-Wh.
Pr-dr	Lab.	-	10	-	-	-	-	-
Rx-dr	Driving	~4.2 kms	7 - 9	35 (Max. 40)	18 - 22	0 - 6 / 10 m ²	0 - 3 / 10 m ²	0 - 2 / 20 m ²
By-dr	Driving	~5.5 kms	7 - 10	40-45	15 - 25	6 - 10 / 10 m ²	3 - 6 / 10 m ²	2 - 4 / 20 m ²
Rt-dr	Driving	~2.5 kms	3 - 4	30 - 35	10 - 15	0 - 4 / 10 m ²	0 - 2 / 10 m ²	0 - 1 / 20 m ²
Po-dr	Lab.	-	5	-	-	-	-	-

Legends: Pr-dr: Pre-driving; Rx-dr: Relax Driving; By-dr: Busy Driving; Rt-dr: Return Driving; Po-dr: Post Driving; Lab.: Laboratory; min: Minutes; Kmph: Kilo meters per hour; Ped.: Pedestrian Count; 2-Wh.: Two Wheeler Count; Bi.: Bicycle Count; 4-Wh.: Four Wheeler Count

2.4. Annotation of Scenarios based Data

Inferring from the driving scenarios and routes followed for data collection we annotated the data in two categories (a) affective state data and (b) stress-trends data. We sought feedback from drivers and based on their perceptions about the routes and the expected level of stress experienced by them, assigned affective state data into four different stress classes and annotated them as Level-1, Level-2, Level-3 and Level-4. The Level-1 stress level here refers to the data collected during pr-dr and po-dr driving scenarios. The other three classes have been shown as underlined numbers (2, 3 and 4) in the Figure 2 depending on the likelihood of a particular affective state level. The routes with minimum pedestrian density and driving effort were given a scale of 2

whereas a scale of 3 was given to the routes with a slightly higher traffic and pedestrians. Since the scenarios belonged to a semi-urban setup, we assigned a scale of 4 for the routes with busy driving scenario. Due to the semi-urban demography, the Level-5 class of Very High stress level was not considered, which the drivers would have experienced while driving on a very busy highway with voluminous traffic for longer duration or typically in a metropolitan city. The stress-trend data has been annotated according to the event markers encountered during drive only and has been presented in Table 2 alongwith their weight scores as well as in Figure 2 at appropriate locations with their abbreviated names.

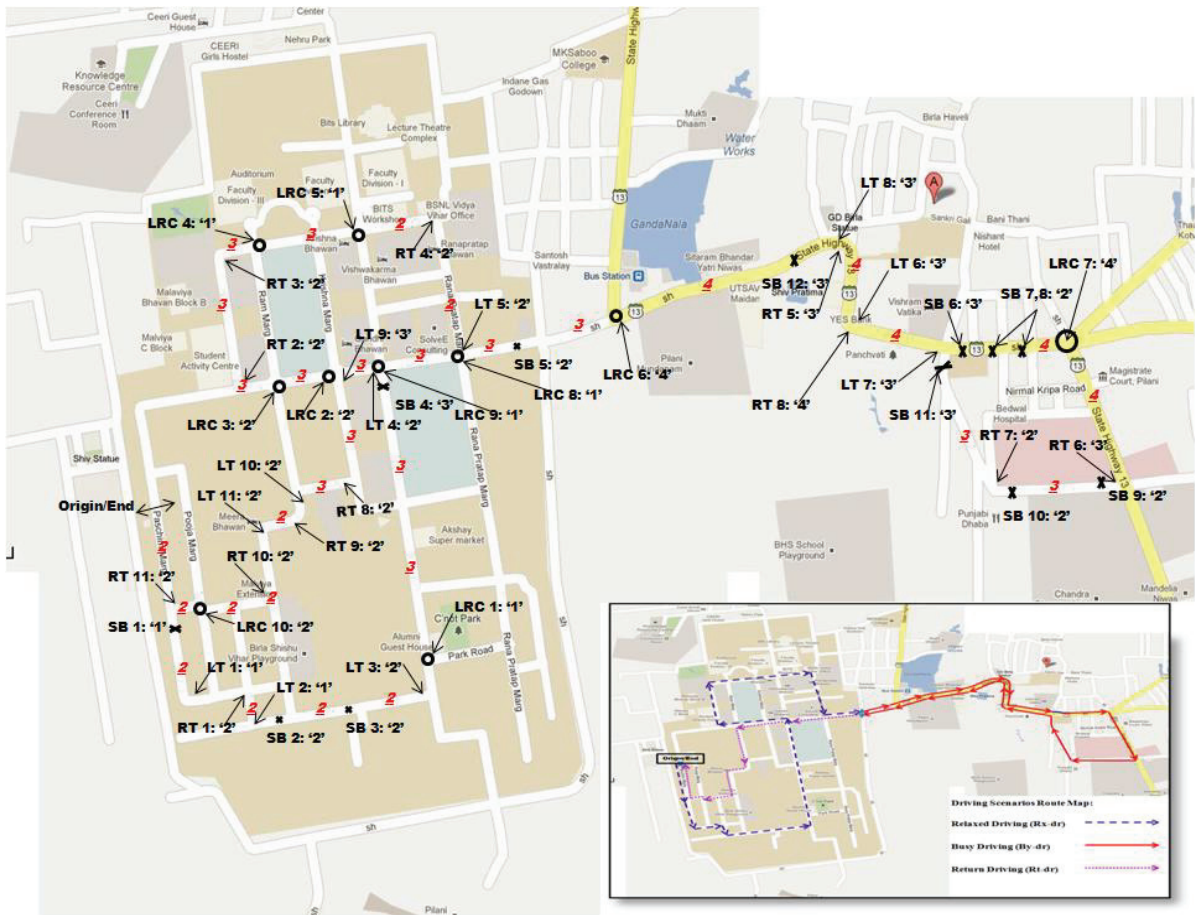


Figure 2. Driving Scenarios Route Map.

Table 2. Stress-Trend Markers and their Weights (Inset: Road route map for Rx-dr, By-dr and Rt-dr)

Stress-Trend Markers	Abbreviations	Weight Score
Left Turn	LT	1 = less (low); 2 = more than 1 (medium) and 3 = greater than 2 (slightly high effort)
Right Turn	RT	2 = approx. same as 2 of LT (medium) and 4 = slightly higher effort required than the 3 of LT
Left-to-Right	LRC	1 = less (low); 2 = approx. same as 2 of LT (medium) and 4 = slightly higher effort required than the 3 of LT
Circle		
Speed Breaker	SB	1 = less (low); 2 = more than 1 (medium) and 3 = greater than 2 (slightly high effort)

3. Data Analysis Methodologies

3.1. Signal Processing: Pre-processing and Normalization

GSR and PPG biosignals are often associated with instantaneous body response to startle, sudden shocks and excitatory stimuli like events often experienced during real-time driving. The dynamic nature of the scenario of data acquisition and changes in sensor orientation/tissue interface results in motion artifacts, impulse noise, sensor baseline-drifts (low frequency errors) and sensor noise (high frequency error). These effects are undesirable and have to be corrected for in physiological signals prior to their applicability in feature extraction for describing the subject's stress level. A time window of 10 seconds (320 samples) was chosen and due care was taken to compensate for the overlaps between the current, preceding and succeeding windows. The choice of the window size is motivated by the observation that a typical GSR response waveform has a time span of about 10 seconds. Such a typical waveform is illustrated in Figure 3. The collected signals were median filtered with a filter of order 3 (for GSR) and order 4 (for PPG) for removal of spikes and impulse noise and baseline drift corrected by de-trending.

GSR and PPG signals are influenced by life-style parameters and their base-line DC value during the experiment is subjected to their activity prior to the experiment. However, since incremental changes are often associated with only changes in the signal response and not its absolute value, signal normalization is often advocated as a reliable pre-processing stage prior to feature extraction. Thus, the filtered physiological signals are min-max normalized using the minima and maxima extracted from the 30-60 second signal window of the Pr-dr scenario (Lowest Stress).

3.2. Signal Processing: GSR

In particular, GSR signal is associated with the physiological response controlled by the sympathetic nervous system and is measured as function of sweat gland activity. The GSR signal is decomposed into constituent *tonic* and *phasic* components which correspond to response over an extended duration of time and short duration of time respectively¹⁶. The *tonic component* is extracted by low pass filtering the normalized GSR signal using a Butterworth 3rd Order Low-pass filter with a cutoff frequency of 0.16 Hz. The *phasic component*, on the other hand, is extracted using band-pass filtering using a Butterworth 3rd order filter with a pass-band of 0.16 Hz to 2.1 Hz. The features extracted from each of these decomposed signal components is discussed and tabulated in Table 3. Further, the feature extraction process entails peak and point of onset detection

which are extracted from the GSR signal using a combination of Central Difference 1st Derivative and Ktonas' 7-point Lagrangian 2nd derivative as illustrated by Zhai et al.¹⁸. Signal points corresponding to zero-crossing in the first derivative and minima in the second derivative were considered as possible peak coordinates. The minimum peak-to-peak interval is 0.5 seconds and multiple peaks within this interval were eliminated in favor of the GSR peak with the highest signal amplitude.

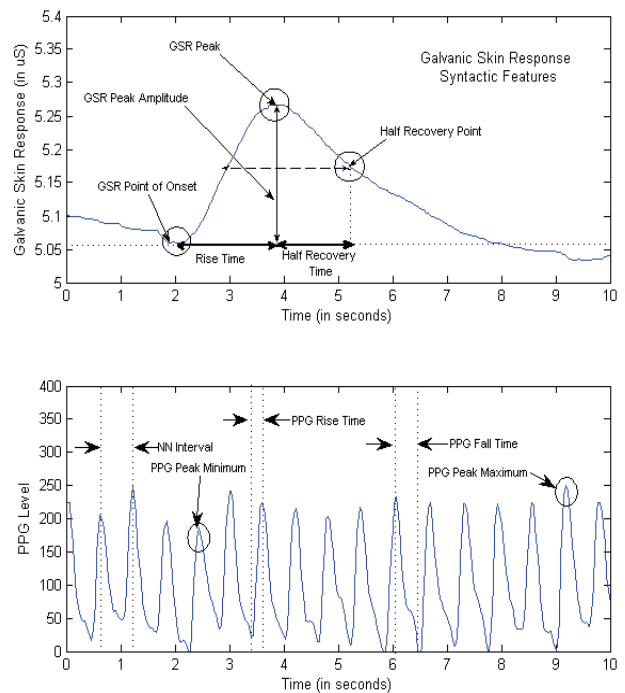


Figure 3. GSR and PPG Syntactic Features during Busy Driving.

3.3. Signal Processing: PPG

For signal analysis associated with PPG, the signal peaks lost due to motion artifacts are corrected using geometric reconstruction of lost peaks by sub-segment replacement using cross-correlation detection method as used by Weng et al.¹⁹. This is followed by peak and trough detection in the reconstructed PPG signal for calculation of the instantaneous heart rate and the corresponding NN interval for feature extraction from HRV (Table 3). The peaks in PPG signal correspond to local maxima and troughs correspond to local minima and these values have been associated with the occurrence of systole and diastole in the heart beat. The minimum peak-to-peak and trough-to-trough interval is 0.5 seconds and peaks and troughs within this interval is eliminated in favor of peak with absolute maximum value and trough with the absolute minimum value. These peaks and troughs were finally utilized to extract instantaneous heart rate (HR) and HRV Linder et al.²⁰. These direct and derived signals were used to extract physiological features as described in the next subsection.

3.4. Feature Extraction and Selection

Since in our study, we are relying on the features derived from the sensed physiological signals due to their clinical significance and established use in identifying stressful events. We extracted features categorized as (a) statistical – characterizes stochastic nature of physiological signals (b) syntactic – derived from geometry of signals and provide contextual structural information and (c) transform based – derived using signal transformation techniques²¹. This resulted in a total of 39 features shown in Table 3 with their mathematical formulae. The syntactic features extracted from GSR and PPG signals have been shown in Figure 3.

The extraction algorithms involved have been discussed in Singh et al.¹¹. The importance of statistical features in physiological signal processing was reported by Lessard²². The feature extraction methodology and clinical significance of GSR features was investigated by Healey and Picard⁷, Schmidt and Walach¹⁶ and Soleymani et al.²³. Researchers^[19,20,24,25,26] have discussed the importance of

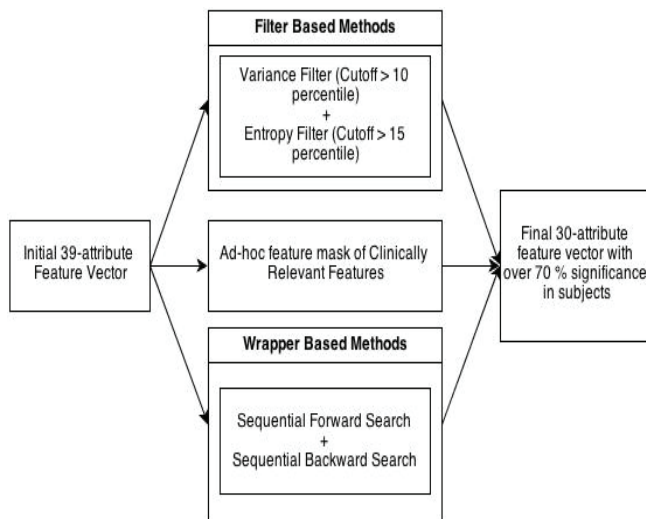


Figure 4. Feature Selection Techniques Adopted.

PPG features and reported about their suitability in such applications. Certain clinically significant parameters like pulse, heart rate (HR) and heart rate variability (HRV) etc. have been successfully extracted by^[13,20] from the morphology of the pulsatile component of PPG signal. These features were found by^[18,27,28] to be as reliable parameters, useful in mental health monitoring of a subject. After these feature extraction routines, each signal segment produced an array of 39 features which constitute the feature vector for that particular time segment.

To select attributes from the feature set that are critical and relevant to represent the problem in a regression or classification model, we adopted the techniques depicted in Figure 4. The feature vector was passed to a variance filter and an entropy filter separately, their combination is known as filter based feature selection. The cutoff for feature

selection was fixed at greater than 15 percentile for entropy filter and 10 percentile for variance filter. The output of these two filters were ANDed to obtain a new feature set resulting in total 26 features whose score were equal and above 70%.

The wrapper based feature selection method is a combination of sequential forward selection (SFS) and sequential backward selection (SBS) where features are added and removed in sequence to meet an objective criterion respectively. The SFS and SBS also resulted in a number of features which were again compared with the 26 features computed from filter based approach and finally we were able to select 27 features which had a score of again 70% and above. Finally, we added an ad-hoc feature mask based on our literature survey to ensure that no clinically significant feature was lost. We obtained a total of 30 features after ORing the ad-hoc feature mask with that of the filter and wrapper based methods, tabulated in Table 3.

Extracting only the selected features ensures reduced time overhead and optimal usage of computing resources. Further, this hybrid combined approach ensures maximal preservation of information (both clinical and classifier-performance driven) as the removed features are redundant and noisy. Table 3 shows the 30 selected features which were found to be significant after the feature selection routine (in the abbreviations column as italicized and boldface), also the individual features selected by each algorithm is shown in columns A-D, whereas F shows the final selection. The concatenated matrix (feature vector matrix) of these 30 feature vectors was used for further classification.

3.5. Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) are mathematical models based on biological network model of a human brain used to solve pattern classification problems by providing an input pattern as a feature vector to one of the target classes²⁹. In applications where training features exhibit non-linearity and the decision boundaries are modeled as a non-linear function in the feature space, ANNs would be a preferred classifier³⁰. ANNs also works reliably with noisy data and can model both categorical and continuous features^{31,32}. Since it is possible that there may still be some noise present even with required signal processing and due to the non-linear nature of data collection, selection of ANN classifiers is justified. The observations and interactions of the experimenter with the drivers led to model the present stress-classification problem as a multiclass classification problem³³, (Level-1 to Level-4 as discussed in Section 2.4) instead of a binary class of just low or high stress class.

Assessment of Driver Stress from Physiological Signals

Table 3: Extracted Features and Formulae

No.	Feature Name	Abbreviation	Description/Formula	A	B	C	D	F
GSR Statistical Features								
1	GSR Mean ²²	GM	$GM = \frac{1}{N} \sum_{i=1}^N X_i$; where x_i is signal value and N is number of samples	√	√	√	√	√
2	GSR Energy ²²	GE	$GE = \sum_{i=1}^N \frac{x_i^2}{f_s}$; where f_s is signal sampling frequency	√	√	√	√	√
3	GSR Time Duration ²²	GTD	$GTD = \left(\frac{\sum_{i=1}^N (t - t_0)^2 x_i^2}{\sum_{i=1}^N x_i^2} \right)^{1/2} * f_s$	√	x	√	x	√
4	GSR Bandwidth ²²	GB	$GB = \frac{1}{2\pi} \left(\frac{\sum_{i=1}^N (x_i - x_{i-2})^2}{\sum_{i=1}^N x_i^2} \right)^{1/2}$	x	√	√	x	√
5	GSR Time Bandwidth Product ²²	GTBP	$GTBP = GTD * GB$	x	√	x	x	x
6	GSR Dimensionality ²²	GD	$GD = 2 * GTD * GB + 1$	x	√	√	x	√
GSR Syntactic Features								
7	GSR Peak Rise Time Sum ¹⁶	GPRTS	Peak Rise time = Time of Occurrence of Peak - Time of Point of Onset	√	√	√	√	√
8	GSR Peak Amplitude Sum ¹⁶	GPAS	Peak-Amplitude = GSR value at Peak- GSR value at Point of Onset	x	x	x	√	√
9	GSR Peak Energy Sum ⁷	GPES	Peak Energy = 0.5 * Peak Amplitude * Peak Rise Time	x	√	√	√	√
10	GSR Half Recovery Sum ¹⁶	GHRS	Half-Recovery Time = Time of Occurrence of Half Amplitude- Time of occurrence of Peak	x	x	x	x	x
11	GSR First Derivative Average ⁷	GFDA	Average First Derivative= Average of the First Derivative observed in the given segment	x	x	x	√	√
12	GSR Rise Rate Average ²³	GRRRA	Average Rise Rate = Sum Average of 1 st derivative of points with 1 st derivative > Positive Threshold (0.025)	x	x	√	x	√
13	GSR Decay Rate Average ²³	GDRA	Average Decay Rate = Sum Average of 1 st derivative of points with 1 st derivative < Negative Threshold (-0.025)	x	√	x	x	x
14	GSR % Decay ²³	GPD	GSR Percentage Decay = Percentage of Time samples in given segment with 1 st derivative < Zero (0).	√	√	√	x	√
15	GSR No. of Peaks	GNP	Number of peaks in a given segment.	x	√	x	√	√
PPG Syntactic Features								
16	PPG Rise Time ²⁰	PPGRT	Average of (Time of Peak- Time of Preceding Trough) in a segment	√	x	√	√	√
17	Pulse Height Min. ²⁴	PPGPHmin	Maximum and Minimum of (Value of PPG peak- Value of PPG trough) in a segment	√	√	√	√	√
18	Pulse Height Max. ²⁴	PPGPHmax		√	√	√	√	√
19	PPG Fall Time ²⁰	PPGFT	Average of (Time of Trough- Time of Preceding Peak) in a segment	√	√	x	√	√
20	Cardiac Period ²⁰	PPGCP	Average of Period of PPG signal in a segment	x	x	√	x	x
21	Inst. HR ^[25-26]	PPGIHR	60 / (Time Difference between two consecutive peaks)	√	x	√	√	√
HRV Spectral Features derived from PPG								
22	PPG Spectral HR	PPGSHR	60* Frequency maximum in range of 0.5-2.5 Hz in HRV spectrum	√	√	√	√	√
23	Respiration Rate ²⁷	RSP	60* Frequency maximum in range of 0.1-0.25 Hz in HRV spectrum	x	√	√	√	√
24	V. Low Freq. Power ²⁵	VLFP	Power in range of 0.003-0.04 Hz in HRV spectrum	√	x	x	x	x
25	Low Freq. Power ²⁵	LFP	Power in range of 0.04-0.15 Hz in HRV spectrum	√	√	x	√	√
26	High Freq. Power ²⁵	HFP	Power in range of 0.15-0.4 Hz in HRV spectrum	√	√	x	√	√
27	LF/HF Ratio ²⁵	LFHF	LF Power/ HF Power	√	√	√	√	√
HRV Statistical Features derived from PPG								
28	AVNN ²⁸	AVNN	Mean of all NN intervals	√	√	√	√	√
29	SDNN ²⁸	SDNN	Standard deviation of all NN intervals	√	√	√	√	√
30	rMSSD ²⁸	rMSSD	RMS of the sequential differences of the IBI calculated for the whole trial	√	x	√	√	√
31	pNN20 ²⁸	pNN20	% of the number of sequential IBI differences that are over 20 ms	√	x	√	√	√
32	pNN50 ²⁸	pNN50	% of the number of sequential IBI differences that are over 50 ms	√	x	x	x	x
PPG Statistical Features								
33	PPG Mean ²²	PPGM	Expression same as GM	√	√	x	√	√
34	PPG Energy ²²	PPGE	Expression same as GE	√	√	√	√	√
35	PPG First Moment ²²	PPGFM	PPG Mean (PPGM) about the origin	√	x	x	x	x
36	PPG Time Duration ²²	PPGTD	Expression same as GT	√	√	√	x	√
37	PPG Bandwidth ²²	PPGB	Expression same as GB	x	x	x	x	x
38	PPG Time Bandwidth Product ²²	PPGTBP	$PPGTBP = PPGTD * PPGB$	√	x	x	x	x
39	PPG Dimensionality ²²	PPGD	$PPGD = 2 * PPGTD * PPGB + 1$	√	√	x	x	√

Legend: GSR- Galvanic Skin Response; PPG- Photoplethysmography; IBI- Inter Beat Interval; A: Filter based method; B: SFS; C: SBS; D: Literature; F: Final Selection; √: Selected; x: Not Selected; The abbreviations shown as italicized indicate the features which were actually selected after the feature selection algorithm.

The stress-classification neural network consists of an input layer, which takes in an input feature vector representing the affective state of drivers, a hidden layer to model the non-linearities in data and an output layer to predict the target affective state class³⁴. Such networks employ a connectionist approach to compute the interconnection weights and bias parameters which produce the most optimal configuration³⁵.

3.5.1. Neural Network (NN) Classifier Configurations Evaluated

The classification of the annotated data into pre-defined affective states i.e. Level-1 to Level-4 is achieved using the six different NN classifiers as explained in Table 4. Besides the affective state, the stress-trends will have cumulative effect on stress level of the drivers. Therefore detection of annotated stress-trends with minimal false alarm is necessary. The proposed approach to perform an instant-by-instant tracking of alarmable trends is discussed in Section 3.7. The feature vector is divided in the ratio of 60:20:20 for training, cross-validation and testing respectively for training the NN.

An activation function in a neural network calculates the layer's output from its net input. This output is fed into subsequent layers as input³⁴. The activation function used in the present application is the tan-sigmoid function. For a very high value of x , the node sends maximum excitation i.e. 1.

$$\text{tansig}(n) = \frac{2}{(1 - e^{-2x})} - 1$$

Where x =combined input to node.

The NN configurations enunciated above utilize back propagation technique for learning. This method utilizes the gradient of error criterion with respect to weights for a given input by propagating the error through the networks. The optimization scheme used is the Levenberg-Marquardt Optimization Scheme. It is a variation of gradient search which employs the least squares criterion for optimization. The neural networks were trained with a stopping criterion on 0.05 MSE with no restriction on the number of epochs. The classifier optimality is evaluated using the performance measures described in the following sub-section.

3.5.2. Classifier Performance Measures

The expected output of 4-class classifiers will result in confusion matrices of size 4x4. The confusion matrix of a binary classifier represents the number of true positives (t_p), true negatives (t_n), false positives (f_p), and false negatives (f_n). However for multiclass problems it is necessary to account for the individual class results for efficient interpretations. Therefore to understand the quality of classifications we consider an individual class C_i , where $i = 1, 2, 3, \dots, n$ are the number of classes, and used macro-averaging techniques³³ to define the classifiers performance measures explained in Table 5.

Table 4: Neural Network Classifiers Evaluated

S. N.	Neural Network Classifier	Description
Feed-forward Neural Network: Provide one-way connection from input to output layers and are used in prediction and pattern recognition problems ³⁴ .		
1.	Multi-Layer Perceptron (1-Hidden Layer) Neural Network (MLP1NN)	The MLP has a fully-connected network structure which enables the pattern of activation in a particular layer at each time step to influence its behavior in the next time step. We envisaged that there might be some inherent overlap between medium, moderate and high affective states data of drivers. Thus, the classifier should be capable of handling non-linearities. The MLP has been found to be capable of being trained in a highly non-linear separable classification problems like the present scenario provided sufficient number of neurons in the hidden layer ³⁴ .
2.	Cascade Forward Backpropagation Neural Network (CASFNN)	CASFNN has connections from the input layer to every previous layer with other characteristics similar to MLP. These networks are trained faster because each neuron is trained independently but may suffer from over fitting problems when the training data used is noisy ³⁶ . The choice of this network for evaluation was made to achieve faster training time with acceptable classification rate.
3.	Feed Forward Time-Delay Neural Network (TDNN)	TDNNs are also similar to MLPs but the inputs to a node also contain some previous time steps realized using tapped-delay lines besides the immediate outputs of previous nodes. These networks can learn precise weight patterns from imprecisely prepared training data ³⁷ and trains faster because the tapped delay line appears only at the input without any feedback loops. These networks were chosen because they are suitable for time series data prediction, which in this case are the selected features representing the present affective state of the driver. We evaluated two separate configurations of this network FFTD-D1NN and FFTD-D2NN with two separate time delays d1 and d2.
Dynamic or Recurrent Neural Network: Useful in time-series prediction and non-linear dynamic problems. They have dynamic neurons, memory and recurrent feedback connections making them suitable as context-aware NNs.		
4.	Elman Back-Propagation Neural Networks (ELMNN)	Elman Networks have a recurrent layer with feed-forward structure. Recurrence simplifies the learning process by allowing the networks to remember states from the past. The context layer copies the output of the layer and uses it as an extra input signal in the next time step in addition to the hidden layer ³⁸ . In the present context it is necessary to check how the reduced complexity in the design affects the predictive ability. Therefore ELMNN was selected.
5.	Layer Recurrent Neural Networks (LRNN)	LRNNs are similar to ELMNNs but each recurrent layer has a tapped delay associated with it. LRNN is useful in evaluating the instances correctly because the outputs depend not only on the current inputs but also on previous states due to the feedback available which enables moving window analysis ³⁹ . The recurrency facilitates the design of a dynamically stable stress classifier due to the presence of delay elements to model the latencies observed in stimuli during the driving process.

Table 5: Classifier Performance Measures

Evaluation Metric	Formula	Importance
Precision	$Precision_{avg} = \left(\sum_i^l \frac{t_{p_i}}{t_{p_i} + f_{p_i}} \right) / l$	Measure of exactness or fidelity ³³ (correctly identified instances of a relevant subset)
Sensitivity	$Sensitivity_{avg} = \left(\sum_i^l \frac{t_{p_i}}{t_{p_i} + f_{n_i}} \right) / l$	The ability of a test to correctly identify positive results ³³
Specificity	$Specificity_{avg} = \left(\sum_i^l \frac{t_{n_i}}{t_{n_i} + f_{p_i}} \right) / l$	The ability of a test to correctly identify negative results ³³
Accuracy	$Accuracy_{avg} = \left(\sum_i^l \frac{t_{p_i} + t_{n_i}}{t_{p_i} + f_{p_i} + t_{n_i} + f_{n_i}} \right) / l$	Overall classification accuracy ³³
Area Under the ROC Curve	$AUC_{avg} = \sum_i^l AUC(i, rest_i) / l$	Trade-off parameter between sensitivity and specificity. AUC value range between 0 and 1.0 ⁴⁰ . For multiclass classification, averaged AUC is computed by considering the one-against-all configuration ⁴¹ (i.e. a c-dimensional classifier as c 2-dimensional classifiers).
Kappa Statistics	$k = \frac{P(a) - P(e)}{1 - P(e)}$ <p>Where: P(a) = Relative observed agreement among the classes P(e) = Probability that agreement is due to chance.</p>	Measure of inter-observer reliability. kappa coefficient ⁴² : ≤ 0 = poor, 0.01 - 0.20 = slight, 0.21 - 0.40 = fair, 0.41 - 0.60 = moderate, 0.61 - 0.80 = substantial, and 0.81 - 1.0 = almost perfect.

3.6. Affective State Classification

While designing an affective state NN classifier, it is important to select a proper window size for optimal performance⁴³. Another important aspect in designing such networks is selection of the number of hidden neurons. The "fixed" approach suggested by Kaastra and Boyd⁴⁴ is best suitable for offline computation despite being time consuming, where a group of networks are trained using different number of neurons. The network satisfying the least error criterion is selected finally. It should be also noted that to avoid overfitting and since we wanted to compare a set of NN classifiers we stuck to just a single hidden layer⁴⁴.

The concatenated matrix of feature vectors extracted earlier alongwith the target vector (representing the stress-classes) were fed to the classifiers selected. The output of the classifiers was expected to be a confusion matrix representing interrelationships between the bio-signal features and the driver's affective state in the form of a stress-class. The real challenge while analyzing the results for such a multiclass classification problem lies in selecting (a) a proper window size (b) the number of neurons in the hidden layer and (c) an optimum classifier which best recognizes the given classes out of this data which is dependent on several parameters. To incorporate this, the training protocol for each NN classifier involved (a) varying non-overlapping window sizes (5 seconds - 30 seconds) (b) varying number of neurons (5 - 30) and (c) training the six selected network configurations and extracting evaluation metrics for each combination. Besides these requirements the classifiers should consistently provide a solution for a large population among the selected drivers. Therefore, a two-fold analysis was performed to meet this requirement. Out of the total 20 drivers, data from 14 drivers were used

for single-turn analysis. This is because during the data collection experiment, for 6 drivers either the physiological data was lost due to the machine errors or misinterpreted by the experimenter and also in some cases the stress-trend markers were not recorded which is important for correlating between affective state and stress-trend detection phases.

3.6.1. Single-Turn Analysis

The single-turn analysis aims at identifying the optimal network configuration with the least inter-observer variability and consistent performance. This method is proposed if the driver assist system designer aims to adopt a universal classifier to be ported into the system's inference engine. The single-turn analysis was performed using the following steps:

- obtained the average of classifier performance measures (e.g. precision, sensitivity and specificity etc.) for each window size (5 to 30) against the six neurons values (5 to 30) selected for each of the six classifiers.
- obtained the individual desirability⁴⁵ for each of the classifier performance measures using their averages obtained, for maximizing a response using the formula shown in Eqn. (1).
- the individual desirability so obtained were used to find the optimum window size, number of neurons in hidden layer and the classifier. Further, the individual desirabilities obtained for each window size were used to obtain an *overall desirability* Eq. (2). The maximum of the individual desirabilities is tabulated in Table 7.

$$d_i = \begin{cases} 0 & ; \text{if } \hat{Y} < Y_{i(\min)} \\ \left(\frac{\hat{Y}_i - Y_{i(\min)}}{Y_{i(\max)} - Y_{i(\min)}} \right)^r & ; \text{if } Y_{i(\min)} < \hat{Y} < Y_{i(\max)} \\ 1 & ; \text{if } \hat{Y} > Y_{i(\max)} \end{cases} \quad (1)$$

$$\text{Overall Desirability} = \sqrt[3]{\left(\frac{\text{Precision}_{\text{avg}} - \text{Precision}_{\min}}{\text{Precision}_{\max} - \text{Precision}_{\min}} \right) \times \left(\frac{\text{Sensitivity}_{\text{avg}} - \text{Sensitivity}_{\min}}{\text{Sensitivity}_{\max} - \text{Sensitivity}_{\min}} \right) \times \left(\frac{\text{Specificity}_{\text{avg}} - \text{Specificity}_{\min}}{\text{Specificity}_{\max} - \text{Specificity}_{\min}} \right)} \quad (2)$$

where

\hat{Y} = each individual response

d_i = individual desirability

$Y_{i(\max)}$ = highest value of Y_i

$Y_{i(\min)}$ = minimum acceptable value of Y_i

r = user defined value ($r = 1$, desirability increases linearly)

3.6.2. Multi - Turn Analysis

Multi-turn analysis was performed to assess whether drivers will have similar effect on their stress levels with same set of experimental protocol and scenarios if they drive multiple times. This analysis aims to identify the most optimal NN configuration for a particular driver which has the least intra-observer variability and high reliability. This analysis approach is to be adopted if the driver assist system designer decides to port driver-specific individualized classifiers into the system's inference engine. Physiological data was collected from six drivers depending upon their availability. This varied data acquisition states and times ensured the required variability which is necessary for an unbiased analysis. In this case since the number of turns each driver completed varied, desirability function approach can not be applied to get the optimum classifier. Therefore, individual feature vector for the respective multiple drives was trained as discussed earlier. The average values for four cardinal classifier performance measures (precision, sensitivity, specificity and classification accuracy) were computed for each window size (5 to 30) against the six neurons values (5 to 30) for each of the six classifier configurations. To find a best fit among the classifiers used for affective state monitoring, we have compared variables such as window size and the number of neurons in the hidden layer against the classifier evaluation parameters.

3.7. Stress-Trend Analysis

During real-time driving, changes in physiological signal base level, morphology, frequency characteristics and statistical parameters are often attributed to instantaneous reflexes and stimuli caused due to *stress-trends*. Real-time detection of such stressful events would help the machine understand the context of the driver's environment and assess the effect of these stimuli on the accumulation of driver stress. Typical stress-trends observed in the driving route has been shown in Figure 2, whereas in Table 2 the annotation of stress-trend markers and the rationale behind the weighing methodology has been tabulated.

For applications which require real-time monitoring of variations in signal's characteristics, Triggs' Tracking Variable (TTV) approach has been a very popular choice^{11,46,47}. The TTV value is a signal detection index and is calculated using the difference between the actual value of a feature and the value predicted using the exponential weighted moving average of the previous values. The TTV value can be used to determine and track *stress-trends* in a feed-forward manner through online monitoring of variations in the observed signals as the absolute value indicates the significance of the change observed. TTV takes values between -1 to +1 which corresponds to 100% certainty of decreasing and increasing trends in the tracked feature⁴⁸. The mathematical description of TTV calculation in the form of pseudocode algorithm is presented in Table 6.

Selection of the exponential smoothing constant ' α ' is crucial which determines the time constant for exponential weighting. Calculation of ' α ' is dependent upon the number of observations ' n ' in a segment using the relationship $\alpha = 2/(n+1)$. A large value of ' α ' is more sensitive to the changes in patterns while estimating the current value of observations. A tracking vector comprising of the TTV values of individual features is generated⁴⁹ for 10 different values of ' α '. Modulating the ' α ' parameter, helps track the changes in the physiological signal at different time-scales. Smaller the ' α ', larger the time window considered for analyzing the incremental changes. These individual tracking vectors adaptively track the incremental changes in the feature values and is used as the input vector to train the neural network classifiers selected for *stress-trend* detection. We evaluated four neural network classifiers using the similar approach followed for affective state detection by using six different numbers of neurons (5 - 30) in the hidden layers for all ' α ' values. The NNs were trained as a regression problem using a stopping criterion of the mean square error value (MSE) of 0.0.

4. Results

The results obtained from both the single-turn and multi-turn analysis for affective state detection and for the stress-trend detection has been discussed individually in the following subsections.

Table 6. Algorithm Pseudocode for TTV calculation

Algorithm $ttv := \text{Triggs' Tracking Variable}$ **Inputs** $D_t :=$ array of feature values observed $\alpha :=$ smoothing constant; between 0.0 and 1.0 which determines the time constant for exponential weighting.**Output** $ttv :=$ Triggs' tracking variable

Begin:

Initialization $u_{t-1} := v_{av}$; exponentially weighted avg. of 1st monitoring segment $s_{t-1} := v_{av}/100$; initial error in prediction for the 1st segment $mad_{t-1} := v_{av}/10$; initial mean absolute deviation**Main Routine**S1: $u_t := \alpha d_t + (1-\alpha)u_{t-1}$; predicted value for present segmentS2: $e_t := d_t - u_{t-1}$; error in predictionS3: $s_t := \alpha e_t + (1-\alpha)s_{t-1}$; smoothened errorS4: $mad_t := \alpha |e_t| + (1-\alpha)mad_{t-1}$; mean absolute deviationS5 $ttv_t := s_t / mad_t$; Triggs' tracking variable for the tth segment**Updation** $u_{t-1} := u_t;$ $s_{t-1} := s_t;$ $mad_{t-1} := mad_t;$ **End****4.1. Results: Single-Turn Affective State Analysis**

The output of the classifiers resulted in confusion matrices of size 4x4 for each window size and the neurons, which was used to calculate the classification performance measures for a particular NN configuration. It can be noticed from Table 7 that to get an optimum window size we can get a trade-off by considering all the variables and evaluation parameters. The degree of suitability of a classifier depends on its predictive ability (precision), identification of true positive rates (sensitivity) and true negative rates (specificity) which is satisfied in this case with a reasonable desirability criterion while maximizing a response. Although a higher percentage of specificity may sound unwarranted but it may be noted that identification of false alarm is as much important as the identification of true

alarms. Hence we have two options (i) CASFNN (window size, WS = 25 and no. of neurons, N = 25) and (ii) MLP1NN (WS = 30 and N = 25). This selection is vindicated from the boxplot diagram (Figure 5) of performance measures for the selected classifier configurations from Table 7. The interquartile ranges (IQR) for CASFNN (WS = 25 and N = 25) configuration is symmetric about the median for almost all performance measures except for the AUC. Additionally the median for this configuration is also above all the configurations. Whereas for the MLP1NN (WS = 30 and N = 25) configuration has inconsistency due to the variable IQRs and presence of outliers.

Table 7. Optimum Window Size Selection for Single Turn Drives

Window Size (Seconds)	Overall Desirability	No. of Neurons in Hidden Layer	Optimum Classifier	Mean Precision	Mean Sensitivity	Mean Specificity
5	0.99957	30	CASFNN	71.20%	69.80%	91.35%
10	0.96078	30	FFTD-D1	72.58%	70.69%	91.90%
15	0.98911	25	CASFNN	72.94%	72.59%	92.10%
20	0.98248	30	MLP1NN	75.85%	72.83%	92.10%
25	0.98455	25	CASFNN	77.94%	78.20%	93.73%
30	0.97329	25	MLP1NN	76.01%	75.38%	92.96%

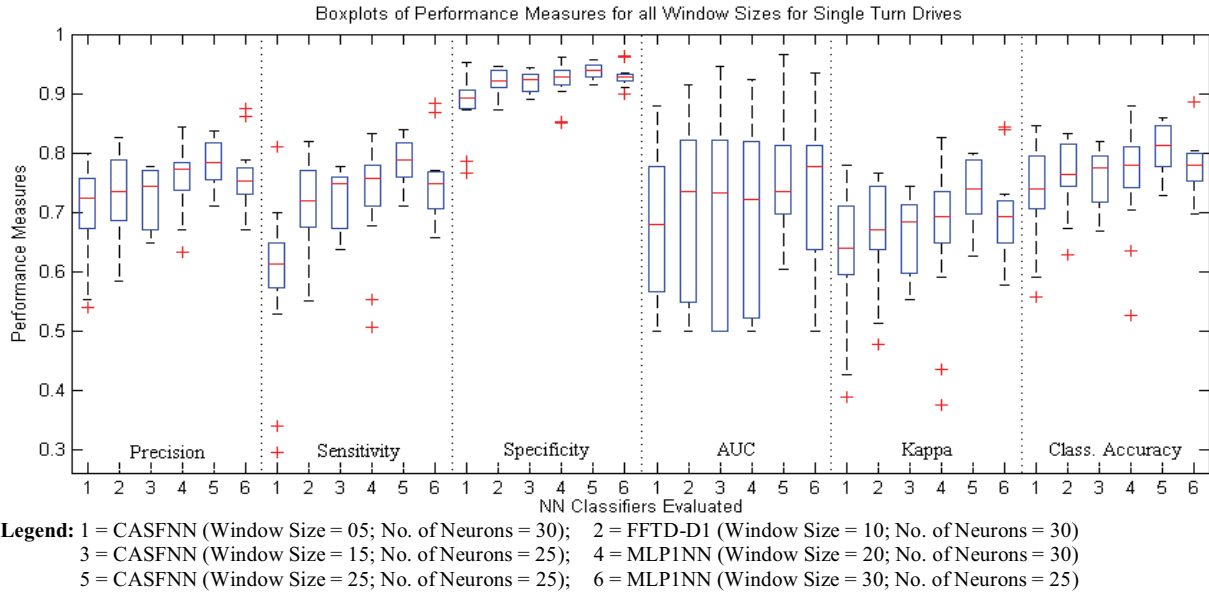


Figure 5: Boxplots of Performance Measures for Single Turn Drives

4.2. Results: Multi-Turn Affective State Analysis

Table 8 presents the results obtained by considering the individual averages for six drivers according to the number of turns completed. The maximum value of performance measures for each drivers is highlighted. Although, the results obtained can not be termed as unique but the observations from the multi-turn analysis can be summarized as below:

- the window size can be fixed between 15 to 30 with 25 being most suitable.
- the number of neurons in hidden layers can be fixed as either 25 or 30.

- the CASFNN classifier performs better as compared to others with MLP1NN being the second choice.
- the effect of number of turns have a little effect on the results obtained as almost all the performance measures have standard deviation very close to their means.
- the classification accuracy is close to 80% which was also obtained for single-drive for the configuration CASFNN (WS = 25; N = 25).

Considering only the overall cardinal performance measures it is clear that an appropriately configured and trained CASFNN classifier could be better choice in the proposed application.

Table 8. Multi-Turn Analysis considering Individual Averages

Drivers (No. of Turns)	Performance Measures	CASFNN (WS=5; N=30)		FTDD1NN (WS=10; N=30)		CASFNN (WS=15; N=25)		MLP1NN (WS=20; N=30)		CASFNN (WS=25; N=25)		MLP1NN (WS=30; N=25)	
		Avg.	STD	Avg.	STD	Avg.	STD	Avg.	STD	Avg.	STD	Avg.	STD
D1 (5)	Precision	68.21%	7.80	68.12%	9.70	65.81%	12.94	73.90%	7.72	75.10%	8.48	69.87%	16.65
	Sensitivity	64.86%	9.28	66.64%	8.99	63.55%	15.86	71.61%	8.87	72.86%	12.86	69.61%	16.86
	Specificity	90.25%	2.89	91.50%	2.47	89.57%	4.99	92.22%	2.41	92.24%	4.09	91.48%	5.31
	Class. Accu.	76.80%	6.04	74.40%	7.36	75.50%	8.36	79.98%	7.26	78.38%	10.44	76.30%	15.99
D2 (5)	Precision	74.44%	11.51	72.27%	6.25	74.95%	9.29	81.52%	2.60	76.02%	2.86	74.62%	10.84
	Sensitivity	73.04%	10.73	69.03%	9.25	74.46%	9.55	81.19%	2.93	76.38%	2.93	73.96%	12.50
	Specificity	92.27%	3.42	91.47%	2.36	92.87%	2.69	94.57%	0.80	93.19%	0.84	92.16%	3.84
	Class. Accu.	76.82%	9.91	74.88%	5.60	78.30%	7.84	83.25%	2.33	78.64%	2.50	76.17%	11.34
D3 (5)	Precision	71.71%	5.06	71.89%	3.57	77.63%	3.15	74.17%	5.18	74.93%	3.38	76.15%	3.96
	Sensitivity	71.26%	5.37	70.65%	3.53	77.64%	3.71	73.18%	7.14	74.30%	3.55	74.40%	2.96
	Specificity	91.77%	1.62	92.06%	0.75	93.40%	1.03	92.42%	2.04	92.74%	1.30	93.05%	0.94
	Class. Accu.	75.03%	5.06	75.89%	2.42	80.01%	2.67	77.52%	5.17	77.86%	4.08	78.61%	2.67
D4 (4)	Precision	71.38%	7.59	71.89%	6.22	75.35%	3.02	78.66%	0.48	78.99%	4.77	74.06%	6.25
	Sensitivity	69.20%	6.96	71.63%	7.24	75.63%	1.70	76.33%	1.19	77.82%	3.87	74.45%	7.27
	Specificity	93.17%	3.60	92.05%	1.74	92.76%	0.32	93.03%	0.64	93.76%	1.11	92.99%	1.18
	Class. Accu.	80.01%	12.18	75.97%	5.04	77.69%	1.12	79.79%	1.82	81.36%	3.37	78.16%	3.83
D5 (3)	Precision	63.13%	9.66	74.19%	3.84	74.49%	6.87	75.74%	2.94	68.71%	22.52	77.52%	14.29
	Sensitivity	62.79%	9.31	72.72%	4.34	72.77%	8.83	75.17%	2.69	66.94%	26.33	78.32%	14.24
	Specificity	89.20%	3.27	92.48%	1.53	91.81%	2.66	92.68%	0.95	89.85%	8.60	95.42%	4.53
	Class. Accu.	67.28%	9.44	76.87%	4.22	75.30%	8.08	77.48%	3.08	68.57%	26.37	86.01%	14.25
D6 (2)	Precision	59.83%	9.29	63.14%	1.08	64.87%	15.79	73.56%	9.05	79.02%	0.12	77.16%	5.68
	Sensitivity	59.01%	8.15	57.89%	6.75	65.32%	15.05	73.52%	8.15	79.66%	0.49	73.63%	5.09
	Specificity	87.93%	3.28	87.26%	2.96	90.66%	3.89	92.39%	2.83	94.05%	0.06	92.56%	2.17
	Class. Accu.	63.79%	10.81	65.07%	4.77	71.55%	11.43	76.70%	8.78	81.58%	0.20	78.43%	6.62

Table 9. Optimum Classifier for Stress-Trend Detection

S.N.	Alpha	CASFNN MSE (R-Value)	MLP1NN MSE (R-Value)	FFTD-D1NN MSE (R-Value)	FFTD-D2NN MSE (R-Value)	No. of Neurons	Optimum Classifier
1.	0.6670	0.005505 (0.997998)	0.011999 (0.995686)	0.007465 (0.997007)	0.0084 (0.996921)	25	CASFNN
2.	0.5000	0.012149 (0.995332)	0.00812 (0.996949)	0.012331 (0.995857)	0.010361 (0.996349)	30	MLP1NN
3.	0.4000	0.005247 (0.997968)	0.016963 (0.993788)	0.011403 (0.995845)	0.008055 (0.996903)	25	CASFNN
4.	0.3330	0.015521 (0.994667)	0.009856 (0.996518)	0.022055 (0.991536)	0.008408 (0.997135)	25	FFTD-D2NN
5.	0.2850	0.009346 (0.996387)	0.019854 (0.992524)	0.014407 (0.994964)	0.017403 (0.994163)	20	CASFNN
6.	0.2500	0.024556 (0.989857)	0.008048 (0.996968)	0.015601 (0.994842)	0.010163 (0.996148)	25	MLP1NN
7.	0.2220	0.005128 (0.998126)	0.01084 (0.996142)	0.015407 (0.994937)	0.012306 (0.996271)	30	CASFNN
8.	0.2000	0.008923 (0.99697)	0.016781 (0.993525)	0.016469 (0.994497)	0.019968 (0.992459)	30	CASFNN
9.	0.1818	0.012568 (0.995154)	0.018661 (0.992761)	0.020861 (0.99242)	0.010468 (0.99607)	30	FFTD-D2NN
10.	0.1667	0.011979 (0.995846)	0.01306 (0.994865)	0.024377 (0.990805)	0.02079 (0.992207)	25	CASFNN

4.3. Results: Stress-Trend Analysis

It was observed that for all the classifiers the MSE values settled close to '0' for neurons greater than 15 and the CASFNN classifier has produced a trained data set with minimum MSE errors. However to obtain an optimum classifier, we compared the average values for MSE and R-square for all the 14 drivers for each ' α ' and observed that in both the cases the CASFNN classifier gave best MSE and R-square values tabulated in the Table 9. The number of neurons in hidden layers can be either 25 or 30 for the configuration selected. The value of smoothing constant (α) can be used as the tuning parameter for stress-trend analysis. Each ' α ' value is inversely related to the time required to determine the stress-trends. In a very dynamic scenario ' α ' has to be tuned higher (like in high stress scenario as inferred from affective state detection) and in a low stress scenario it has to be tuned down. Therefore it can be concluded that the CASFNN classifier alongwith a suitable value of ' α ' will give best results.

5. Discussion and Concluding Remarks

Design and development of a safety-critical wearable driver assist system needs an intelligent inference framework for assessment of stress experienced by the drivers under real-time driving scenarios. This requires continuous monitoring of drivers' affective state, detecting on-road stress-trends and cumulatively accounting for increasing stress level. Physiological signals viz. GSR and PPG were collected using a wearable signal acquisition unit. The driving scenarios were chosen to emulate a semi-urban driving setting with variability in the road structure, traffic composition and pedestrian density throughout the drive.

For the presented work, data was acquired and annotated from 20 drivers and a total of 38 driving turns, however 14

driver's data was utilized for single-turn and stress-trend analysis whereas 6 driver's data (6+18 = 24) was analyzed for multi-turns. The overall methodology adopted in this work is summarized in the Figure 6.

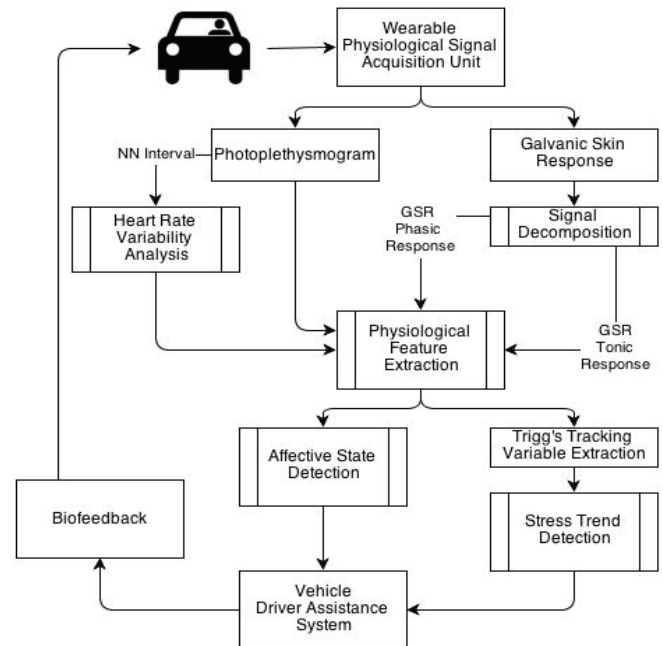


Figure 6. Driver Stress Level Analysis: The Complete Flow.

A two-fold analysis approach viz. Single-Driving Turn Analysis and Multiple-Driving Turn Analysis was adopted to evaluate the classifier performance for affective state detection. These tested not only the accuracy of the classifier and inter-observer performance variability (single-turn analysis) but also evaluated its potential to produce reproducible and reliable results i.e. intra-observer performance variability (multi-turn analysis). The single-turn and multi-turn analyses are aimed at a universal design

and a driver-specific individualized design approaches respectively. The single-turn analysis established that the CASFNN configuration with 25 hidden neurons and a selected window size of 25 seconds performed most optimally minimizing the inter-observer variability with an overall accuracy of 80.61% with a standard deviation of 4.15. The multi-turn analysis test for intra-observer variability reaffirmed that CASFNN configuration with the same configuration is the most optimal network configuration for designing individualized classifiers as well.

To monitor instantaneous reflexes and stimuli experienced by the driver contributing to stress, we adopted a novel stress-trend analysis approach using Trigg's tracking variable. This stress-trend analysis was modeled as a NN regression problem and the observed results further established CASFNN configuration as the most consistent network with the least mean square error and maximal R-value of over 95% for each of the identified stress trends. Porting and testing of this algorithm on a body-worn wearable computer can enable us to activate rescue and relaxation procedures in case of accidents and alarmable situations.

The robustness of the developed algorithm will be further enhanced by training the WDAS using data collected from a larger representative population, by considering more hazard prone scenarios and longer real-time driving environments. For real-time testing and evaluation, this phase of work would be followed by porting of these algorithms on a wearable embedded test bed. Complexity minimization followed by resource optimization and ergonomic design would lead to a more economical and reliable compute-infrastructure for future wearable driver assist systems.

6. Acknowledgment

The author wishes to thank Dr. Stephen P. Linder of Dartmouth College, Hanover, USA for his technical support in morphological analysis of PPG signals. We also want to thank Dr. B. K. Rout at BITS Pilani for his valuable suggestions. One of us (R.R.Singh) likes to acknowledge the timely help rendered by Jitin, Vamsi, Srikumar and Tejaswi of BITS Pilani.

References

1. *Global status report on road safety: time for action*. Geneva, World Health Organization, 2009 (www.who.int/violence_injury_prevention/road_safety_status/2009).
2. Traffic Safety Facts 2008: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System, *National Highway Traffic Safety Administration (N.H.T.S.A.)*, Report No. DOT HS 811 170, (US Department of Transportation, Washington DC, 2008).
3. C. Liu and T. J. Ye, Run-Off-Road Crashes: An On-Scene Perspective, *National Highway Traffic Safety Administration, Report No. DOT HS 811 500*, (US Department of Transportation, Washington DC, July 2011).
4. L. James and D. Nahl, Dealing with stress and pressure in the vehicle. Taxonomy of driving behavior: Affective, cognitive, sensorimotor, Chapter 3 in *Driving lessons: Exploring systems that make traffic safer*, J. Peter Rothe, (Ed.). (Edmonton, Canada: University of Alberta Press, 2003), pp. 21-50.
5. C. L. Lisetti, and F. Nasoz, Using noninvasive wearable computers to recognize human emotions from physiological signals, *EURASIP Jour. Of Applied Signal Processing*, **11**, 2004, pp. 1672–1687.
6. C. D. Katsis, N. Katertsidis, G. Ganiastas and D. I. Fotiadis, Toward Emotion Recognition in Car-Racing Drivers: A Biosignal Processing Approach, *IEEE Trans. On Systems Man And Cybernetics – Part A: Systems and Humans*, **38** (3), 2008, pp.502–512.
7. J.A. Healey and R. W. Picard, Detecting Stress During Real-World Driving Tasks Using Physiological Sensors, *IEEE Trans. on Intelligent Transportation Systems*, **6** (2), 2005, pp. 156-166.
8. R. Banerjee, From Research to Classroom: A Course in Pervasive Computing, *IEEE Pervasive Computing*, **4** (3), 2005, pp. 83-86.
9. M. Patel, S. K. Lal, D. Kavanagh and P. Rossiter, Applying neural network analysis on heart rate variability data to assess driver fatigue, *Expert Systems with Applications*, **38** (6), 2011, pp. 7235–7242.
10. A. Riener, A. Ferscha and M. Aly, Heart on the road: HRV analysis for monitoring a driver's affective state, in *Proc. 1st Int. Conf. Automotive User Interfaces and Interactive Vehicular Applications*, (AutomotiveUI, 2009), pp. 99-106.
11. R. R. Singh, S. Conjeti and R. Banerjee, An Approach for Real-Time Stress-Trend Detection in Physiological Signals in Wearable Computing Systems for Automotive Drivers, in *Proc. 14th Int. IEEE Annual Conf. on Intelligent Transportation Systems*, (Washington DC, USA, 2011), pp. 1477 – 1482.
12. Dinesh Mohan, Road Accidents in India, *IATSS Research*, **33** (1), 2009, pp. 75-79.
13. H. H. Asada, P. Shaltis, A. Reisner, S. Rhee and R.C. Hutchinson, Mobile Monitoring with Wearable Photoplethysmographic Biosensors, *IEEE Engg. in Med. and Bio. Magazine (Special Issue on Wearable Sensors / Systems and Their Impact on Biomedical Engineering)*, **22**, 2003, pp. 28-40.
14. Nexus-10 Wireless Monitoring and Biofeedback System. [Online]. Available: <http://www.mindmedia.nl/english/nexus10.php>.
15. E. Kaniusas, Biomedical Signals and Sensors I: Linking Physiological Phenomena and Biosignals, Ch. 1, (Springer, Heidelberg, 2012).
16. S. Schmidt, and H. Walach, Electrodermal Activity (EDA) – State-of-the-Art Measurement and Techniques for Parapsychological Purposes, *The Jour. of Parapsychology*, **64**, 2000, pp. 139-163.
17. R. R. Singh and R. Banerjee, Multi-parametric Analysis of Sensory Data collected from Automotive Drivers for Building a Safety-Critical Wearable Computing System, in *Proc. 2nd Int. Conf. on Computer Engg. and Tech.*, (Chengdu, China, 2010), pp. VI-355 -360.

18. J. Zhai, A. B. Barreto, C. Chin, and C. Li, Realization of Stress Detection using Psychophysiological Signals for Improvement of Human-Computer Interactions, in *Proc. of the SoutheastCon*, 2005, pp. 415-420.
19. J. Weng, Z. Ye, and J. Weng, An Improved Pre processing Approach for Photoplethysmographic Signal, in *Proc. IEEE Engg. in Med. and Biology 27th Annual Conf.*, (Shanghai, China, 2005), pp. 41-44.
20. S. P. Linder, S. M. Wendelken, E. Wei, and S. P. McGrath, Using the Morphology of Photoplethysmogram Peaks to Detect Changes in Posture, *Jour. of Clinical Monitoring and Computing*, **20** (3), 2006, pp. 151-158.
21. E. J. Ciaccio, S.M. Dunn and M. Akay, Biosignal Pattern Recognition And Interpretation Systems, Part 2 of 4: Methods for Feature Extraction and Selection, *IEEE Engg. in Med. and Biology*, 1993, pp. 101-112.
22. C. S. Lessard. Signal Processing of Random Physiological Signals, Ch. 3, (Morgan & Claypool, San Rafael, California, 2006).
23. M. Soleymani, G. Chanel, J. J. M. Kierkels and T. Pun, Affective Ranking of Movie Scenes Using Physiological Signals and Content Analysis, in *Proc. 2nd ACM workshop on Multimedia semantics*, 2008, pp. 32-39.
24. M. Shamir, L. A. Eidelman, Y. Floman, L. Kaplan, and R. Pizov, Pulse oximetry plethysmographic waveform during changes in blood volume, *British Jour. of Anaesthesia*, **82** (2), 1999, pp. 178-181.
25. N. Hjortskov, D. Risse'n, A. K. Blangsted, N. Fallentin, U. Lundberg, K. Søgaard, The effect of mental stress on heart rate variability and blood pressure during computer work, *European Jour. Applied Physiology*, **92**, 2004, pp. 84-89.
26. D. W. Ryoo, Y. S. Kim, and J. W. Lee, Wearable Systems for Service based on Physiological Signals, in *Proc. IEEE Engg. in Med. and Biology 27th Annual Conf.*, (Shanghai, China, 2005), pp. 2438-2440.
27. D. L. Partin, M. F. Sultan, C. M. Thrush, R. Prieto, and S. J. Wagner, Monitoring Driver Physiological Parameters for Improved Safety, in *Proc. 2006 SAE World Congress*, (Detroit, Michigan, 2006), ISBN: 0-7680-1633-9.
28. D. Giakoumis, A. Vogianou, I. Kosunen, K. Moustakas, D. Tzovaras, and G. Hassapis, Identifying Psychophysiological Correlates of Boredom and Negative Mood Induced During HCI, in *Proc. 3rd Int. Joint Conf. on Biomedical Engg. Sys. and Tech. (BIOSTEC)*, (Valencia, Spain, 2010), pp. 3-12.
29. Anil K. Jain, J. Mao and K. M. Mohiuddin, Artificial neural networks: A tutorial, *IEEE computer*, **29** (3), 1996, 31-44.
30. J. A. K. Suykens, J. P. L. Vandewalle, and B. L. R. De Moor, *Artificial Neural Networks for Modeling and Control of Non-Linear Systems*. (Springer, 1995).
31. A. B. M. S. Ali, S. A. Wasimi, *Data Mining: Methods and Techniques*, Chapter 5, (Cengage Learning, India Edition, New Delhi, 2009).
32. R. O. Duda, P.E. Hart, and D. G. Stork, *Pattern Classification*, Chapter 6, (Wiley India, Second Edition, New Delhi, 2006).
33. M. Sokolova, and G. Lapalme, A systematic analysis of performance measures for classification tasks, *Information Processing and Management*, **45**, 2009, pp. 427-437.
34. S. Haykin, *Neural Networks*, Chapters 3 - 6, (Pearson Education, Second Edition, New Delhi, 2001).
35. R. Schalkoff. *Pattern Recognition: Statistical, Structural and Neural Approaches*, Chapters 10 - 13, (Wiley India, New Delhi, 2005).
36. V. Schetinin, A Learning Algorithm for Evolving Cascade Neural Networks. *Neural Processing Letter*, **17** (1), (Kluwer, 2003), pp. 21-31.
37. K. J. Lang and A. H. Waibel, A Time-Delay Neural Network Architecture for Isolated Word Recognition, *Neural Networks*, **3**(1), 1990, pp. 23 - 43 .
38. A. Grüning, Elman Backpropagation as Reinforcement for Simple Recurrent Networks. *Neural Computation*, **19** (11), 2007, pp. 3108-3131.
39. Q. Liu, and J. Wang, A One-layer Recurrent Neural Network for Convex Programming, in *Proc. IEEE Int. Joint Conf. on Neural Networks*, 2008, pp. 83-90.
40. T. Fawcett, An Introduction to ROC Analysis, *Pattern Recognition Letters*, **27**, 2006, pp. 861-874.
41. C. Ferri, J. Hern'andez-Orallo, R. Modroiu, An Experimental Comparison of Performance Measures of Classification, *Pattern Recognition Letters*, **30** (1), 2009, 27 - 38.
42. J. R. Landis and G. G. Koch, The Measurement of Observer Agreement for Categorical Data, *Biometrics*, **33**, 1977, pp. 159-74.
43. R. J. Frank, N. Davey, and S. P. Hunt, Time Series Prediction and Neural Networks, *Journal. Intelligent Robotics Systems*, **31** (1-3), 2001, pp. 91-103.
44. I. Kaastra, and M. Boyd, Designing a Neural Network for Forecasting Financial and Economic Time Series. *Neurocomputing*, **10**, 1996, pp. 215 - 236.
45. G. Derringer and R. Suich, Simultaneous Optimization of Several Response Variables, *Journal of Quality Technology*, **12** (4), 1980, pp. 214-219.
46. W. W. Melek, Z. Lu, A. Kapps, and W. D. Fraser, Comparison of Trend Detection Algorithms in the Analysis of Physiological Time-Series Data, *IEEE Trans. On Biomedical Engineering*, **52** (4), 2005, pp. 639-651.
47. P. Yang, *Adaptive Trend Change Detection and Pattern Recognition in Physiological Monitoring*. Doctoral Thesis, (University Of British Columbia, 2009).
48. G.S. Cembrowski, J. O. Westgard, A. A. Eggert and E. C. Toren, Trend Detection in Control Data: Optimization and Interpretation of Trigg's Technique for Trend Analysis, *Clinical Chemistry*, **21** (10), 1975, pp. 1396-1405.
49. C. E. Hope, C. D. Lewis, I. R. Perry, and A. Gamble, Computed trend analysis in automated patient monitoring systems, *British Jour. of Anesthesia*, **45**, 1973, pp. 440-448.