IJEE Volume-5 • Number-2 December 2013 pp. 13-20 ISSN: 0973-7383

A Novel Method of Stress Detection using Physiological Measurements of Automobile Drivers

Mandeep Singh, Abdullah Bin Queyam
Department of Electrical & Instrumentation Engineering, Thapar University, Patiala, INDIA
mandy_tiet@yahoo.com, abdullahbinqueyam@gmail.com

Abstract
Stress while driving is an important factor in many number of fatal road accidents worldwide. There has been much work done in driver stress detection. In this research, we present a method based on a correlation analysis and developed a mathematical function for the estimation of automobile driver stress level. The proposed methodology monitors driver's stress level using features extracted from selected physiological parameters. The results obtained indicate a strong correlation between the stress level of driver and the stress function formed. Threshold approach is used to perform a classification of affective states as "Low Stress", "Moderate Stress" and "High Stress" based on different traffic conditions. The stress function acts as a direct indicator of stress level of the automobile diver whose physiological parameters are monitored continuously under variable traffic conditions.
Keywords: Automobile driver, stress, correlation, function, physiological parameters, features extraction, classification.

1. Introduction

Approximately 1.24 million people die every year on the world's roads, and another 20 to 50 million sustain nonfatal injuries as a result of road traffic crashes. Road traffic injuries are estimated to be the eighth leading cause of death globally, with an impact similar to that caused by many communicable diseases [1]. Alertness of a driver while driving an automobile is an important parameter to be monitored and maintained in order to avoid the vehicular accidents. Measuring physiological conditions offers a practical method of determining the mental stress level of an automobile driver. Driver's stress level can be measured using recording acquired from physiological sensors [2]. Such recordings can be used by stress detection system installed within the automobiles. These in-vehicle electronic systems may further improve the decision making capability of the driver, and makes an intelligent transportation system.

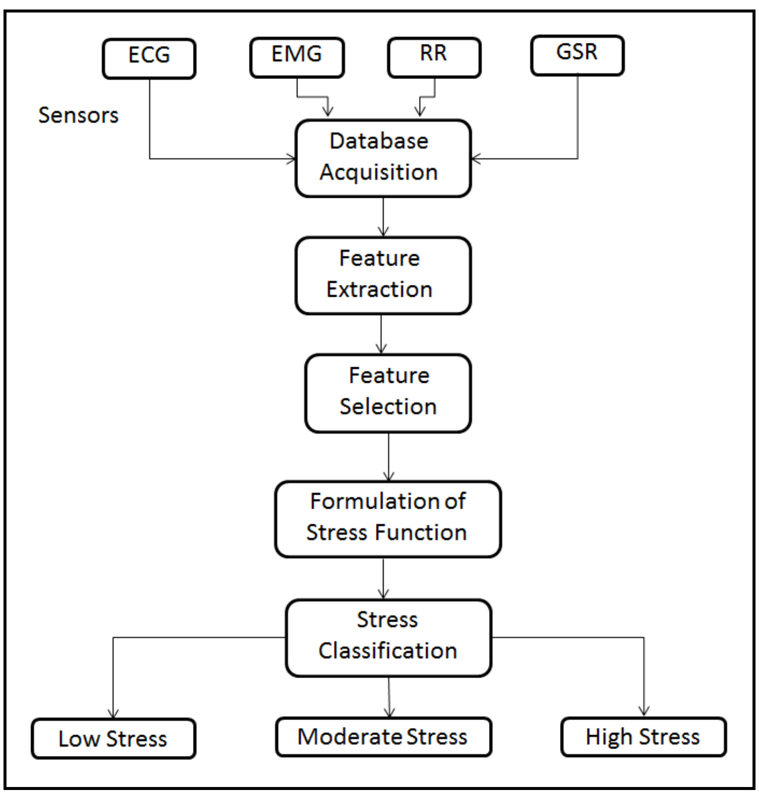

Fig. 1. General Layout of the proposed Stress Detection System

In our previous work we have validated that Galvanic Skin Response (GSR) and Heart Rate (HR) are the two most important physiological parameters for detecting driver's drowsiness, fatigue or stress level under variable traffic conditions. Several other physiological parameters like Respiration Rate (RR), ElectroMyoGram (EMG) also affect the driver stress level by a small factor [3]. Their effect can also be combined to form a reliable and more accurate stress detection system/model. The work described in this research aims to design a stress detection system that can monitor the stress level of a driver using physiological parameters such as GSR, EMG, RR and HR as shown in Fig. 1. We acquired raw physiological database from PHYSIONET website [4].

Healey and Picard presented methods to evaluate driver's relative stress level using physiological signals. The study indicated that Galvanic Skin Response (GSR) and heart rate (HR) signals were significantly correlated to driver stress [5]. Stress is known to activate the Sympathetic Nervous System (SNS) [6]. Much research has already been done on the detection of stress from physiological parameters that are influenced by the SNS. Examples are muscle activity, heart rate, heart rate variability, skin conductance and pupil diameter [7–9]. Other studies have shown that a combination of these physiological parameters facilitates differentiation between stressful situations and situations without stress [10–16].

The data acquired from Healey and Picard's [5] experiment lacks the information regarding the duration of each Rest, City and Highway driving task, but the same durations were mentioned in Ahmet Akbas [17] research and we found it to be very useful in our research for stress detection in automobile drivers.

2. Feature Selection

As many pattern recognition techniques were originally not designed to cope with large amounts of irrelevant features, combining them with Feature Selection techniques has become a necessity in many applications. The objectives of feature selection are manifold, the most important ones being:

a) to avoid overfitting and improve model performance, i.e. prediction performance in the case of supervised classification and better cluster detection in the case of clustering,
b) to provide faster and more cost-effective models, and
c) to gain a deeper insight into the underlying processes that generated the data.

For selecting the best features out of 8 statistical features we used software called 'WEKA' [18]. WEKA (Waikato Environment for Knowledge Analysis) is a data mining system developed by the University of Waikato in New Zealand that implements data mining algorithms using the JAVA language. WEKA is a state-of-the-art facility for developing machine learning (ML) techniques and their application to real-world data mining problems. It is a collection of machine learning algorithms for data mining tasks. The algorithms are applied directly to a dataset. WEKA is open source software issued under General Public License. WEKA contains tools for data pre-processing, classification, regression, clustering, association rules, attribute selection and visualization. It is also well-suited for developing new machine learning schemes.

Features are specified attribute numbers before applying attribute selection analysis on our database as shown in Table 1.

Table 1
Attributes used in 'WEKA' shown by corresponding attribute no.

Attribute No.	Feature/Attribute Name
1	Mean hand GSR
2	Mean foot GSR
3	Mean EMG
4	rms EMG
5	RR
6	Mean HR
7	HRV
8	NHRV

After applying feature selection method the results obtained shows that attribute no. 1, 2, 3, 4, 5, 6 are the attributes selected by different algorithms as shown in Table 2.

Table 2
Results obtained after using different feature selection algorithms.

S No.	Drive No.	Attribute Selector Used	Selected Attributes	Overall Selected Attributes
1	All Drive	ClassifierSubsetEval	1,2,3,4,5,6	1,2,3,4,5,6
2	All Drive	ChiSquaredAttributeEval	1,2,3,4,6	
3	All Drive	CfsSubsetEval	1,2,3,4,5,6	
4	All Drive	SVMAttributeEval	1,2,3,5,6	
5	All Drive	WrapperSubsetEval	1,2,3,5,6	

Out of 8 statistical features we selected only 6 features as shown in Table 3 for our further research.

Table 3
Final attributes selected using different algorithms in 'WEKA' Software

Attribute No.	Feature/Attribute Name
1	Mean Hand GSR
2	Mean Foot GSR
3	Mean EMG
4	rms EMG
5	RR
6	Mean HR

Applying correlation analysis on the 6 selected statistical features, we observed that in most of the cases the correlation between Mean Hand GSR vs Mean Foot GSR and Mean EMG vs rms EMG are very strong as shown in Table 4. Thus, using Mean EMG instead of rms EMG will not make much difference in our aim to form a function that will act as a stress level indicator.

Table 4
Correlation coefficients obtained between Mean Hand GSR vs Mean Foot GSR and Mean EMG vs rms EMG for all ten Drives'.

S No.	Driver's data set No.	Corresponding *r* values between	
		Mean Hand GSR vs Mean Foot GSR	Mean EMG vs rms EMG
1	Drive 05	0.97	0.97
2	Drive 06	0.47	0.97
3	Drive 07	0.82	0.81
4	Drive 08	0.92	0.97
5	Drive 09	0.96	0.91
6	Drive 10	0.84	0.96
7	Drive 11	0.88	0.96
8	Drive 12	0.24	0.99
9	Drive 15	0.80	0.94
10	Drive 16	0.79	0.97

Also shown in Table 4 is correlation between Mean Hand GSR and Mean Foot GSR. Theoretically, there should be high correlation between these two physiological parameters. In case of some drives like Drive06 and Drive12, the low correlation may be due to some data acquisition problems. These drives thus may need special techniques for stress classification. This is elaborated in the later sections of this paper.

3. Formation of Stress Detection Function

In statistics, regression analysis is a statistical technique for estimating the relationships among variables. More specifically, regression analysis helps one understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed. Regression analysis is also used to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships. Our approach in classification of stress level is to consider the rest period as low stress condition, highway driving as moderate stress condition and city drive as high stress condition. Neglecting all other factor towards the contribution of stress, an analysis of correlation of various physiological parameters with the three mentioned states of mental stress has been made in this paper. Further we have chosen numerical value 1 for low stress, 3 for moderate stress and 5 for high stress with gradual transitions in between [3].

We applied Regression Analysis by taking Stress Class of automobile driver as Output variable and Mean Hand GSR, Mean EMG, RR, Mean HR and Mean Foot GSR as Input variables as shown in Table 5.

Table 5
Variables used in Regression analysis

Inputs (X)	Output (Y)
1. Mean Hand GSR ($\mathbf{X_1}$) 2. Mean EMG ($\mathbf{X_2}$) 3. RR ($\mathbf{X_3}$) 4. Mean HR ($\mathbf{X_4}$) 5. Mean Foot GSR ($\mathbf{X_5}$)	Stress Class (**LS, MS and HS**)

We have divided our analysis into 4 Cases:

1. **Case 1** includes Stress Class as Output variable and Mean Hand GSR, Mean EMG, RR and Mean HR as Input variables.
2. **Case 2** includes Stress Class as Output variable and Mean Hand GSR, RR and Mean HR as Input variables.
3. **Case 3** includes Stress Class as Output variable and Mean Foot GSR, RR and Mean HR as Input variables.
4. **Case 4** includes Stress Class as Output variable and Mean Hand GSR, Mean Foot GSR, RR and Mean HR as Input variables.

The results were computed on MS Excel and the coefficients of function obtained from Case 1 are shown in Table 6. Similarly, Case 2, Case 3 and Case 4 have been analysed and corresponding functions are formed as shown in Table 7.

Table 6
Results obtained from Regression Analysis of Case 1.

S No.	Driver's data set No.	Coefficients				Intercept
		Mean Hand GSR (X_1)	Mean EMG (X_2)	RR (X_3)	Mean HR (X_4)	
1	Drive 05	0.6707	0.3482	0.2471	0.4868	3.3000
2	Drive 06	0.1276	0.1737	0.7271	0.6647	3.1500
3	Drive 07	0.9273	0.2820	0.3728	-0.3789	3.2500
4	Drive 08	1.0022	-0.0488	0.5814	0.4096	3.2000
5	Drive 09	1.2903	-0.1174	0.1910	0.1554	3.6286
6	Drive 10	0.3559	0.4445	0.7108	0.0614	3.0750
7	Drive 11	1.4506	-0.1217	0.0900	0.1633	3.1000
8	Drive 12	1.0239	-0.0340	-0.0845	0.1104	3.0750
9	Drive 15	0.9763	0.1877	0.5500	-0.0535	3.0260
10	Drive 16	0.9580	-0.3615	0.2602	0.5659	3.5938
MEAN		**0.8783**	**0.0753**	**0.3646**	**0.2185**	**3.2398**

The results obtained in Table 6, shows that Mean EMG has very less impact on change in value of stress function ($\square_1$). Thus, Mean EMG has been neglected in Case 2 to find the function ($\square_2$). We observed that there is a negligible change in value of correlation between Stress Class and function ($\square_2$) with respect to value of correlation between Stress Class and function ($\square_1$). Functions ($\square_3$ & $\square_4$) obtained in Case 3 and Case 4 also show very strong correlation with Stress Class. Table 8 shows a comparison of the correlations obtained between all the four functions formed with Stress Class.

Table 7

Function obtained from Regression Analysis of Case 1, Case 2, Case 3, and Case 4.

Case No.	Function Formed
Case 1	**Function ($\square_1$) = 0.8783(X_1) + 0.0753(X_2) + 0.3646(X_3) + 0.2185(X_4) + 3.2398**
Case 2	**Function ($\square_2$) = 0.8783(X_1) + 0.3646(X_3) + 0.2185(X_4) + 3.2398**
Case 3	**Function ($\square_3$) = 0.7907(X_5) + 0.4184(X_3) + 0.3587(X_4) + 3.2398**
Case 4	**Function ($\square_4$) = 0.4916(X_1) + 0.5257(X_5) +0.2851(X_3) + 0.2007(X_4) + 3.2398**

Table 8

Comparison of Correlation of Stress Class with Functions for All Drives

S No.	Driver's data set No.	Correlation Coefficients				
		Stress Class vs Mean Hand GSR	Case 1	Case 2	Case 3	Case 4
			Stress Class vs Function ($\square_1$)	Stress Class vs Function ($\square_2$)	Stress Class vs Function ($\square_3$)	Stress Class vs Function ($\square_4$)
1	Drive 5	0.813	0.860	0.853	0.862	**0.862**
2	Drive 6	0.256	0.593	0.582	**0.704**	0.636
3	Drive 7	**0.680**	0.677	0.675	0.602	0.642
4	Drive 8	0.783	0.857	0.858	**0.887**	0.885
5	Drive 9	0.856	0.850	0.853	0.830	**0.860**
6	Drive 10	0.359	0.550	0.540	**0.731**	0.643
7	Drive 11	**0.877**	0.852	0.854	0.776	0.840
8	Drive 12	**0.609**	0.572	0.573	0.471	0.584
9	Drive 15	0.736	0.774	0.770	0.772	**0.802**
10	Drive 16	0.782	0.838	0.846	0.850	**0.864**
Mean Correlation.		**0.6751**	**0.7423**	**0.7404**	**0.7485**	**0.7618**
Score		**3**	**0**	**0**	**3**	**4**

Our previous research showed that stress is most strongly related to the Mean Hand GSR feature of the automobile driver. There are several other physiological features like Mean EMG, RR, Mean HR, Mean Foot GSR that are also affected by the dynamic stress level of the automobile drivers [3]. Thus we have formulated a stress function ($\square$) which combines important features drawn from different sensors in specific proportion. The function ($\square$) obtained have higher degree of strength of correlation with Stress Class as compared to just Mean Hand GSR feature. Changing the features used for the formulation of function ($\square$), changes value of the correlation coefficient. For each drive the best correlation function has been highlighted in Table 8. The score evaluated in the last row of this table is the count of the best correlations.

Function ($\square_4$) has the best correlation results with the stress class as indicated by the overall Mean Correlation and Score value. Thus, instead of Mean Hand GSR, the function ($\square_4$) can directly be used for the detection of dynamic stress level of an automobile driver continuously over a period of time. In our further research of stress classification we have used only function ($\square_4$) as a standard function for determination of stress level.

$$\textbf{Function } (\square_4) = 0.4916(X_1) + 0.5257(X_5) + 0.2851(X_3) + 0.2007(X_4) +$$

4. Stress Classification

Classification is a task of training a model that maps each attribute set X to one of the predefined class labels Y. In our case class labels Y is the Stress Class of automobile driver defined by different traffic conditions (Resting, Highway and City) and attribute set X is the function ($\square_4$) which is a linear combination of four statistical features namely Mean Hand GSR, Mean Foot GSR, Respiration Rate and Mean HR extracted from the physiological parameters. Function ($\square_4$) is the only attribute used for stress classification because this function is alone sufficient for detection of stress level of the automobile driver. Resting phase of the driving segment is considered as Low Stress (LS) category. Highway phase is considered as Medium Stress (MS) category, while driving in City is considered as High Stress (HS) category as shown in Table 9. It is assumed that the change in stress level of the driver is solely due to the variable traffic conditions and all other factor towards the contribution of stress is neglected.

Table 9
Comparison of Correlation of Stress Class with Functions for All Drives

Driving Segments Used	Category	Stress Class
Initial Rest + Final Rest	Relaxed or Low Stress	**LS**
City 1 + City 2 + City 3	High Stress	**HS**
Highway 1 + Highway 2	Moderate Stress	**MS**

In our research the classification of three stress class is done by the use of two threshold values. One threshold value differentiates Low Stress category and Moderate Stress category while other differentiates Moderate Stress category and High Stress category.

E.g. Stress Class LS has min value of 1 and max value of 2, HS has min value of 2 and max value of 3, while HS has min value of 3 and max value of 5. Clearly the two threshold values which classify the stress class into three categories (LS, MS and HS) would be 2 and 3. Similarly, the classifications of all the ten Drives are carried out by using a common threshold value and also by changing the threshold value with respect to individual Drive; results are shown in Table 10.

Table 10
Percentage of Correctly Classified Instances after Classification.

Drive No.	Accuracy for Common Thresholds (3, 3.5)	Accuracy for Individual thresholds	
Drive05	78.7 %	**80 %**	(3.1, 3.6)
Drive06	61.2 %	67.5 %	(3.2, 3.4)
Drive07	68.7 %	68.7 %	(3, 3.5)
Drive08	81.2 %	**83.7 %**	(3, 3.9)
Drive09	70 %	**81.4 %**	(1.9, 3)
Drive10	60 %	66.2 %	(2.6, 3)
Drive11	80 %	**80 %**	(3, 3.5)
Drive12	62.5 %	63.7 %	(2.8, 3.7)
Drive15	83.1 %	**83.1 %**	(3, 3.5)
Drive16	75 %	**82.8 %**	(2.5, 3.2)

It is depicted from Table 10, that the stress function ($\square_4$), alone is able to classify 6 out of 10 drives with accuracy of more than 80% into their respective Stress Class of automobile driver (LS, MS and HS). These accuracies have been highlighted in the table. However, the classification accuracy in case of Drive06, Drive07, Drive10 and Drive 12 is very low. To further improve the accuracy of classification of drives Drive06, Drive07, Drive10 and Drive 12, we have used Artificial Neural Network (ANN) approach. As ANN requires multiple attributes, the attributes used in function ($\square_4$), namely Mean Hand GSR, Mean Foot GSR, Respiration Rate and Mean HR are chosen as four inputs to ANN used for stress classification.

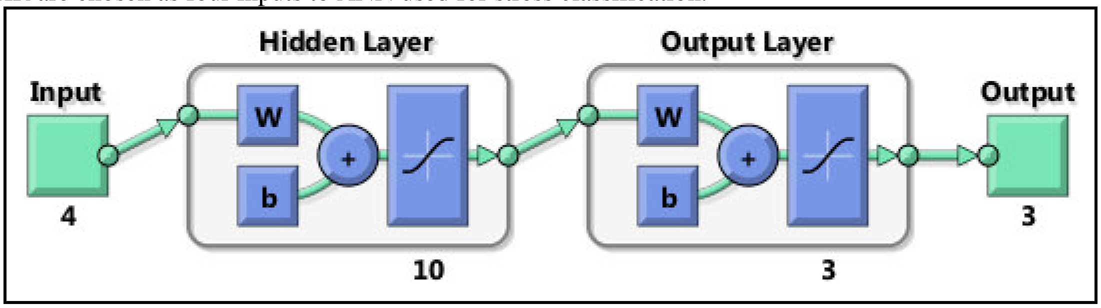

Fig. 2. A two-layer feed-forward network.

A typical two-layer feedforward back propagation network having 4 inputs, 3 outputs, one hidden layer with 10 neurons and a output layer is used. Sigmoid function is used in hidden layer as shown in Fig. 2. The Neural Network Pattern Recognition Tool (nprtool) of MATLAB is employed to create and train a network, and evaluate its performance using mean square error and confusion matrices. The result obtained is shown in Table 11. It clearly illustrates that by using ANN, the percentage of classification accuracy increases by significant amount as compared to the results obtained by using threshold classification.

In summary, a single function ($\square_4$) is sufficient for the stress classification in most of the drives. However, there may be some problems in data acquisition as in Drive06, Drive07, Drive10 and Drive12. This calls for an intelligent approach of using ANN for enhancing accuracy of classification.

Table 11
Percentage of Correctly Classified Instances after using Artificial Neural Network.

Drive No.	Individual threshold for each Drive	Using Artificial Neural Network (feedforward back propagation)	Confusion Matrix
Drive06	67.5 %	85 %	*Predicted Class* $\begin{array}{c} \\ LS \\ MS \end{array}\begin{array}{ccc} LS & MS & HS \\ 27 & 0 & 2 \\ 1 & 11 & 4 \end{array}$ *ual Class*
Drive07	68.7 %	82.5 %	*Predicted Class* $\begin{array}{c} \\ LS \\ MS \end{array}\begin{array}{ccc} LS & MS & HS \\ 24 & 0 & 2 \\ 0 & 13 & 5 \end{array}$ *ual Class*
Drive10	66.2 %	88.75 %	*Predicted Class* $\begin{array}{c} \\ LS \\ MS \end{array}\begin{array}{ccc} LS & MS & HS \\ 29 & 0 & 2 \\ 0 & 12 & 3 \end{array}$ *ual Class*
Drive12	63.7 %	77.5 %	*Predicted Class* $\begin{array}{c} \\ LS \\ MS \end{array}\begin{array}{ccc} LS & MS & HS \\ 26 & 0 & 5 \\ 0 & 6 & 9 \end{array}$ *ual Class*

5. Conclusions

It may be concluded that for threshold based classification, the best suited physiological parameters are Mean Hand GSR, Mean Foot GSR, RR and Mean HR. Since the base physiological parameters are different for different drives, individual thresholds for each drive gives better classification accuracy as compared to the common thresholds. In case of some discrepancies in the data, classification accuracy may fall down. In that case, ANN based classifier is able to give better accuracy.

6. Scope for Future Work

Using the results of this research work, a microcontroller/field-programmable gate array (FPGA) based stress level indicator can be designed on real time monitoring of driver's mental stress. In case the driver is in low or moderate stress, he may be allowed to use navigational tools, else in case of high stress, the driver may be advised only to focus on driving and avoid multitasking, or take rest in between.

References

[1] World Health Organization, "Global Status Report on Road Safety 2013," Geneva, WHO, 2013. Available at: http://www.who.int/iris/bitstream/10665/78256/1/9789241564564_eng.pdf

[2] Mandeep Singh and Abdullah Bin Queyam, "Stress Detection in Automobile Drivers using Physiological Parameters: A Review," *International Journal of Electronics Engineering*, vol. 5, no. 2, Dec 2013.

[3] Mandeep Singh and Abdullah Bin Queyam, "Correlation between Physiological Parameters of Automobile Drivers and Traffic Conditions," *International Journal of Electronics Engineering*, vol. 5, no. 2, Dec 2013.

[4] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215]; 2000 (June 13). PMID: 10851218; doi: 10.1161/01.CIR.101.23.e215
Available at: http://physionet.org/cgi-bin/atm/ATM/

[5] J. A. Healey and R.W. Picard, "Detecting Stress During Real-World Driving Tasks Using Physiological Sensors," *IEEE Trans. Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156-166, June. 2005.

[6] C. Tsigos and G. P. Chrousos, "Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress," *Journal of Psychosomatic Research*, vol. 53, pp. 865–871, 2002.

[7] U. Lundberg, R. Kadefors, B. Melin, G. Palmerud, P. Hassmen, M. Engstrom, and I. E. Dohns, "Psychophysiological stress and EMG activity of the trapezius muscle," *International Journal of Behavioral Medicine*, vol. 1, pp. 354–370, 1994.

[8] M. L. Peters, G. L. R. Godaert, R. E. Ballieux, M. van Vliet, J. J. Willemsen, F. C. G. J. Sweep, and C. J. Heijnen, "Cardiovascular and endocrine responses to experimental stress: Effects of mental effort and controllability," *Psychoneuroendocrinology*, vol. 23, no. 1, pp. 1– 17, Jan. 1998.

[9] J. Zhai and A. Barreto, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables." *In IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 1355–1358, 2006.

[10] I. C. Jeong, D. Lee, S. W. Park, J. I. Ko and H. R. Yoon, "Automobile driver's stress index provision system that utilizes electrocardiogram," *Intelligent Vehicles Symposium, 2007 IEEE* , vol., no., pp. 652-656, 13-15 June 2007.

[11] L. Salahuddin and D. Kim, "Detection of Acute Stress by Heart Rate Variability Using a Prototype Mobile ECG Sensor," *International Conference on Hybrid Information Technology*; vol. 2; pp. 453-459, 2006.

[12] R. R. Singh, S. Conjeti, R. Banerjee, "Biosignal based on-road stress monitoring for automotive drivers," *Communications (NCC), 2012 National Conference on*, vol., no., pp. 1-5, 3-5 Feb. 2012.

[13] G. Wu, G. Liu and M. Hao, "The Analysis of Emotion Recognition from GSR Based on PSO," *Intelligence Information Processing and Trusted Computing (IPTC), 2010 International Symposium on* , vol., no., pp. 360-363, 28-29 Oct. 2010.

[14] J. Wijsman, B. Grundlehner, J. Penders and H. Hermens, "Trapezius muscle EMG as predictor of mental stress," *Wireless Health* 2010 (WH '10). ACM, New York, NY, USA, pp. 155-163, 2010.

[15] J. A. Healey, "Wearable and automotive systems for affect recognition from physiology", MIT Dept. of Electrical Engineering and Computer Science, Ph.D. thesis, 2000.

[16] I. Homma and Y. Masaoka, "Breathing rhythms and emotions," *Experimental Physiology*; 93 (9), pp. 1011-1021, September 1, 2008.

[17] A. Akbas, "Evaluation of the Physiological Data Indicating the Dynamic Stress Level of Drivers," *Scientific Research and Essays*, 6(2): pp.430-439, 2006.

[18] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten; The WEKA Data Mining Software: An Update; *SIGKDD Explorations*, vol. 11, Issue 1, 2009.