

# Feature Selection and Machine Learning Based Multilevel Stress Detection from ECG Signals

Isabelle Bichindaritz<sup>(✉)</sup>, Cassie Breen, Ekaterina Cole, Neha Keshan,  
and Pat Parimi

Computer Science Department,  
State University of New York at Oswego, Oswego, USA  
{ibichind, cbreen, ecole3, neha.keshan,  
patanjali.parimi}@oswego.edu

**Abstract.** Physiological sensor analytics aims at monitoring health as the availability of sensor-enabled portable, wearable, and implantable devices become ubiquitous in the growing Internet of Things (IoT). Physiological multi-sensor studies have been conducted previously to detect stress. In this study, we focus on electrocardiography (ECG) monitoring that can now be performed with minimally invasive wearable patches and sensors, to develop an efficient and robust mechanism for accurate stress identification, for example in automobile drivers. A unique aspect of our research is personalized individual stress analysis including three stress levels: low, medium and high. Using machine learning algorithms from the ECG signals alone, our system achieves up to 100% accuracy and area under ROC curve of 1 depending on the experimental setting in detecting three classes of stress using feature selection from a combination of fiducial points and multiscale entropy as a fine-grained indicator of stress level.

**Keywords:** Machine learning · Sensors · Data mining · Stress medicine · ECG

## 1 Introduction

According to the National Highway Traffic Safety Administration (NHTSA) and the Virginia Tech Transportation Institute (VTTI), lack of attention while driving is found to be the leading cause of automobile accidents in the US in 80% of all crashes and 65% of all near-crashes. These data clearly show that improved attention and close monitoring of drivers' conditions could help increase their safety. Driving in stressful environments such as city or highway prompts for drivers' heightened attention and is also correlated with higher risk of accidents because prolonged stress decreases one's ability to be attentive.

With the availability of portable wearable and implantable devices in the growing Internet of Things (IoT), physiological sensor data analytics will lead to improved health care monitoring [1] and preventive care [2]. Although physiological multi-sensor studies have been conducted with some success to detect stress based on such measures as heart rate variability, skin conductance, respiration rate, electromyogram (EMG), body

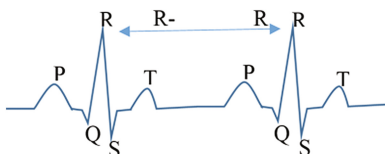
temperature, blood pressure, and electro-encephalogram (EEG). Electrocardiography (ECG) has often been discarded by these studies due to the constraints of the measurements requiring 16 leads and the possible imperfections of the resulting signals, which can fail to detect some heart beats. However, ECG signals are highly valued for the precision of their R-peak detection, leading to excellent heart rate rhythm measurement after preprocessing the signal for missed beats [3]. In addition, ECG monitoring can now be performed with minimally invasive wearable patches and other sensors, which makes stress detection based on them an interesting field of study [4].

In this study, we apply machine learning methods and algorithms to detect stress from ECG signals in subjects under different levels of environmental stress caused by driving conditions. We find that stress levels can be successfully detected from ECG signals alone; with random tree classifier allowing for identification of the three classes of stress, low, medium and high, with up to 100% accuracy depending on the experimental setting, which is a significant improvement on a prior study on the same data set [5]. In particular, classification accuracy was improved by 10% in cross-validation with Multilayer Perceptron.

## 2 Background

The ECG is one of the simplest and oldest cardiac monitors available and yet it can provide a wealth of useful information. ECG represents the electrical activity of the heart muscle as it changes with time [6]. Like other muscles, the cardiac muscle contracts in response to electrical *depolarization* of the muscle cells. It is the sum of this electrical activity, when amplified and recorded for just a few seconds that is known as an ECG.

Important waveforms of an ECG are marked as P, Q, R, S and T (see Fig. 1) and represent the changes in electrical potential as the heart contracts and relaxes. Points P, Q, R, S and T are called fiducial points. Depolarization of the ventricles results in usually the largest part of the ECG signal (because of the greater muscle mass in the ventricles) and this is known as the **QRS complex** [7].



**Fig. 1.** An ECG signal depicting the RR interval

- The Q wave is the first initial downward or ‘negative’ deflection.
- The R wave is the next upward deflection (provided it crosses the isoelectric line and becomes ‘positive’)
- The S wave is the next deflection downwards, provided it crosses the isoelectric line to become briefly negative before returning to the isoelectric baseline.

## 3 Materials and Methods

### 3.1 Data

The ECG signals of stress used in this study were obtained from MIT-BIH PhysioNet Multi-parameter Database [8]. These data sets are part of the experiment conducted by Healey and Picard [3] and has data from 17 participating drivers and eight types of raw data – time stamp, ECG, electromyogram (EMG), foot galvanic skin response (GSR), hand GSR, intermittent heart rate (IHR), marker, and respiration – all acquired from different wearable sensors. During the experiment conducted by Healy and Picard, the drivers drove in Greater Boston area from MIT’s East Garage to River Street Bridge and back through three cities and two highways. The initial rest and final rest states, as well as stress during driving were measured.

The data sets were segmented into three stress levels – low stress (initial rest and final rest), moderate stress (highway), and high stress (cities) – assuming that the stress acquired by subjects is solely based on traffic conditions and for no other reason. The signal classification was carried out by considering the variation in the ECG signals of the three states, low, medium and high, in an individual.

The time durations for each segment – rest, highway, and city - given by Akbas were used to distinguish between the rest, highway and city time periods [9]. The segmentation mark of different driving periods was not clear in seven of the data sets as was also found by Akbas [9]. Consequently, only 10 drivers’ data sets were used for this study. Using the methods available from Physionet [8], an annotation was performed on each data set separately and annotated files were obtained for each driving period of the ten drivers.

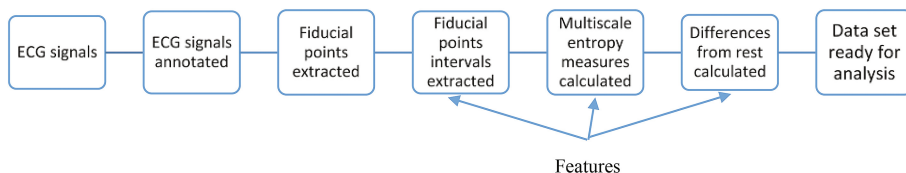
### 3.2 Feature Extraction

Feature extraction was performed to extract 14 different fiducial points (P, Q, R, S) interval features, averaged over the time intervals (see Table 1, left column) from the annotated ECG signals annotations using NetBeans Java platform (see Fig. 2) to produce the required file for classification in Waikato Environment for Knowledge Analysis (Weka) [10]. We have considered all possible signal attributes and their relations (Table 1) in feature extraction to carry out a thorough analysis. We demonstrated in a previous paper [5] that near-perfect classification could be achieved with these 14 features alone – and even a subset of these – for two stress levels. However, results for three stress levels were not convincing, which prompted us to add multiscale entropy to this original set of features.

Therefore, we added variance for these intervals (12 new features), which did not improve the results much. We then performed multiscale entropy analysis [11] of the annotation files. Intuitively, the entropy of a signal measures the amount of disorder and complexity present in this signal. Pathological states have been found to be associated with decreased complexity in signals, and lower multiscale entropy. Examples of pathologies include aging and chronic heart failure. Multiscale entropy consists in extracting from a time series entropy measures associated with several

**Table 1.** Extracted features from the ECG signals

Fiducial features		Entropy features	
Average/Var QRS interval	Average/Var QRS difference	A1 to A20 entropy	A1 to A20 entropy diff.
Average/Var RR interval	Average/Var RR difference		Average entropy diff.
Average/Var QQ interval	Average/Var QQ difference	Entropy variance	Entropy variance diff.
Average/Var SS interval	Average/Var SS difference	Entropy slope 1	Entropy slope 1 diff.
Average/Var QR interval	Average/Var QR difference	Entropy slope 2	Entropy slope 2 diff.
Average/Var RS interval	Average/Var RS difference		
Average beats	Average difference beats		



**Fig. 2.** Feature extraction algorithm flowchart

scales in the signal, according to windows of varying level of granularity, starting from the whole signal and progressively dividing it by a scale factor. The algorithm proceeds in two steps. First for each scale  $i$ , a time series is generated by averaging the samples in windows of length  $i$ . Then the entropy is calculated for each coarse-grained time series by the conditional probability quantifying the likelihood that if two sets of simultaneous data points of a given length have distance  $<r$ , then two sets of simultaneous data points of the given length plus 1 also have distance  $<r$  [11].

With a scale factor of 20, we obtain 20 entropy measures (see Table 2). We can also calculate the average entropy and its variance, as well as the slopes at the beginning of its curve and at its end (see Table 2).

In addition, differences in fiducial point characteristics and multiscale entropy measurements between rest and stress states were recorded (Fig. 2 and Table 2) since the differences from the baseline may be important (as we found in a previous study [5]).

**Table 2.** Accuracy percentage of three classes with 74 attributes

Accuracy (%)	Classification methods					
	Training set	Leave 1 out	2 Folds	10 Folds	75% Split	90% Split
Naïve Bayes	71.64	53.73	50.75	53.73	58.82	85.71
BayesNet	73.13	67.16	62.69	61.19	<b>70.59</b>	85.71
Logistic	<b>100</b>	38.81	53.73	47.76	52.94	57.14
Multi-Perceptron	98.51	50.74	62.69	56.72	58.82	85.71
SMO	85.07	59.70	53.73	58.21	58.82	57.14
IB1	<b>100</b>	41.79	47.76	38.81	52.94	85.71
IBK	<b>100</b>	41.79	47.76	38.81	52.94	85.71
KStar	<b>100</b>	53.73	50.75	53.73	58.82	<b>100</b>
ZeroR	44.78	44.78	44.78	44.78	29.41	41.86
J48	94.03	<b>68.66</b>	59.70	<b>62.69</b>	64.71	71.43
Random Forest	<b>100</b>	61.19	<b>65.67</b>	<b>62.69</b>	58.82	71.43
Random Tree	<b>100</b>	46.27	59.70	56.72	58.82	71.43

### 3.3 Classification

Predicting the level of stress from these ECG signals is a typical classification task in data mining. Three classes are available for classification purposes. Class ‘0’ represents low stress (rest state), class ‘1’ moderate stress (highway driving) and class ‘2’ represents high stress (city driving).

Weka offers many classifiers out of which 12 algorithms from varied types were selected for classification to perform their comparative study [10] (see Table 2).

### 3.4 Assessment

The following six different test and experimental settings were applied on the ECG data sets: Training Set, Leave One Out Cross Validation (LOOC), 2-Folds Cross Validation, 10-Folds Cross Validation, 75% Split, and 90% Split.

The data set analyzed was small since it contained only 68 instances, obtained from the 70 potential instances for 10 drivers and 7 driving intervals for each [9]. Two of these driving intervals could not be analyzed to produce an annotation file with Physionet’s annotators. From these 68 annotated signals, we have removed the data of drive05 highway1 as it is an obvious outlier with an average number of beats per minute of 29 – a highly unlikely figure. Other studies also reported this outlier and removed it. Therefore, the experiments presented below were conducted using the remaining data set of 67 instances with 74 different extracted attributes.

The experiments were conducted on 3 classes of stress – ‘0’ for low stress, ‘1’ for moderate stress, and ‘2’ for high stress. The features extracted consisted in 14 fiducial measures and 48 entropy measures. Each time, classification was performed on all the features because the classification methods used were capable of selecting best features.

74 features were considered (see Table 2):

- 6 average interval durations (QRS, RR, QQ, SS, QR, RS).
- 1 average number of beats per minute.
- 6 average interval durations differences between initial rest and current state (QRS, RR, QQ, SS, QR, RS).
- 1 average number of beats per minute difference between rest and current state.
- 6 variance of interval durations (QRS, RR, QQ, SS, QR, RS).
- 6 variance of interval durations differences between initial rest and current state (QRS, RR, QQ, SS, QR, RS).
- 24 entropy measures (A1 to A20, average, variance, slope between A1 and A2, slope between A19 and A20).
- 24 differences in entropy measures between initial rest and current state (A1 to A20, average, variance, slope between A1 and A2, slope between A19 and A20).

## 4 Results

Accuracy, sensitivity (true positive rate), specificity (true negative rate), and area under ROC curve (AUC) were selected as performance measures. Accuracy was rounded off to two decimal places and the sensitivity, specificity, and AUC were rounded off to three decimal digits. Bold values in the tables represent the best results per column.

### 4.1 Results on 3 Classes and 74 Features

The results presented here are the classification accuracy percentage (Table 2), sensitivity/specificity, and AUC (Table 3) on the data set with three classes, all 74 attributes, and 67 instances.

### 4.2 Results on 3 Classes and 10 Selected Features

The results presented here are the classification accuracy percentage (Table 4), sensitivity/specificity, and AUC (Table 5) on the data set with three classes and 10 selected attributes chosen by automated feature selection in Weka (SVMAttributeEval with Ranker method).

### 4.3 Results Interpretation

The results on 3 classes and 74 features show that the highest accuracy was obtained for J48 (decision tree) on LOOC (68.66%) and Random Forest and J48 on 10-fold cross validation (62.69%). However, with a 90% split, which is acceptable for such a small data set, accuracy reaches 100% for KStar (Table 3). If considering the AUC, Random Forest reaches .832 in 10-fold cross validation and 1 with a 90% split (Table 3).

**Table 3.** Sensitivity/Specificity and AUC of three classes with 74 attributes

Sens/Spec AUC	Classification Methods					
	Training set	Leave 1 out	2 Folds	10 Folds	75% Split	90% Split
Naïve	.716/	.537/	.507/	.537/	.588/786/	.857/
Bayes	.877/.891	.782/.646	.775/.611	.788/.658	.649	.976/1
BayesNet	.731/ .802/.844	.672/ .766/.72	.627/ .768/.741	.612/.75/ .745	<b>.706/</b> <b>.801/.795</b>	.857/ .893/.839
Logistic	<b>1/1/1</b>	.388/ .683/.587	.537/ .791/.7	.478/ .735/.616	.529/ .768/.662	.571/ .679/.667
Multi. Perceptron	.985/ .988/.991	.507/ .747/.718	.627/ .812/.741	.567/ .757/.724	.588/ .786/.718	<b>.857/</b> <b>.893/.976</b>
SMO	.851/ .918/.898	.597/ .782/.749	.537/ .766/.684	.582/ .784/.723	.588/ .786/.716	.571/ .845/.833
IB1	<b>1/1/1</b>	.418/ .673/.546	.476/ .722/.6	.388/ .656/.522	.529/ .761/.645	.857/ .893/.875
IBK	<b>1/1/1</b>	.418/ .673/.546	.476/ .722/.596	.388/ .656/.539	.529/ .761/.645	.857/ .893/.875
KStar	<b>1/1/1</b>	.537/ .748/.717	.507/ .755/.706	.537/ .748/.726	.588/ .786/.736	1/1/1
ZeroR	.448/ .552/0.5	.448/ .552/0.5	.448/ .552/0.5	.448/ .552/.461	.294/ .706/.5	.429/ .571/.5
J48	.94/.971/ .98	.687/ .836/.674	.597/.8/ .679	.627/ .794/.687	.647/ .789/.669	.714/ .869/.72
Random Forest	<b>1/1/1</b>	<b>.612/</b> <b>.755/.818</b>	<b>.657/</b> <b>.803/.799</b>	<b>.657/</b> <b>.761/.832</b>	.588/ .758/.751	.714/ .869/.929
Random Tree	<b>1/1/1</b>	.463/ .723/.593	.597/ .808/.703	.567/ .788/.678	.588/ .833/.71	.714/ .869/.792

The results on 3 classes and 10 automatically selected features (see Table 4) show that the highest accuracy was obtained for J48 on LOOC (73.13%) and Multilayer Perceptron (MLP) on 10-fold cross validation (80.60%). However, with a 90% split, which is acceptable for such a small data set, accuracy reaches 100% for MLP and Logistic (Table 4). Considering the AUC, MLP ranks higher in most categories, including 10-fold cross-validation with .875 and 90% split with 1, also reached by Logistic (Table 5).

Overall, results have significantly improved using feature selection, with or without cross validation. However Random Forest, which performed best on 74 features, due to its ability to discriminate between features, was overtaken by MLP after feature selection.

In our previous studies, the results on 3 classes and 1 feature showed that the highest accuracy was obtained for MLP (neural network) on LOOC (68.66%) and 10-fold cross validation (70.15%), with J48 almost as accurate [5]. By adding variance and multiscale entropy, accuracy has improved by 10% with MLP in cross validation and by 14% in 90% split, which are significant improvements.

**Table 4.** Accuracy percentage of three classes with 10 selected attributes

Accuracy (%)	Classification Methods					
	Training set	Leave 1 out	2 Folds	10 Folds	75% Split	90% Split
Naïve Bayes	73.13	56.72	58.21	58.21	58.82	85.71
BayesNet	68.66	67.16	61.20	65.67	70.59	85.71
Logistic	86.57	71.64	56.72	65.67	41.18	<b>100</b>
Multi-Perceptron	94.03	71.64	<b>73.13</b>	<b>80.60</b>	<b>76.47</b>	<b>100</b>
SMO	77.61	67.16	62.69	70.15	58.82	85.71
IB1	100	64.18	64.18	65.67	64.71	71.43
IBK	100	64.18	64.18	65.67	64.71	71.43
KStar	100	53.73	50.75	55.22	58.82	85.71
ZeroR	44.78	44.78	44.78	44.78	29.41	42.86
J48	86.57	<b>73.13</b>	65.67	67.16	<b>76.47</b>	85.71
Random Forest	100	70.15	64.18	68.66	64.71	85.71
Random Tree	100	67.16	61.19	61.19	58.82	85.71

The best overall results were obtained for the features automatically selected by SVMAttributeEval (see Table 6), which combine the difference between rest and stress for each instance as well as some features independent from the rest state. Selected features include difference of heart rate from rest state, difference of variance in RR interval from rest state, difference of variance in QQ interval from rest state, variance in SS interval, heart rate, difference of variance in SS interval from rest state, difference of variance in entropy A10 and in entropy A7 from rest state, variance in entropy A4, and variance in entropy A13. We see that the addition of multiscale entropy has significantly improved the classification performance with four contributing features.

## 5 Discussion

The particular data set used for the present study was produced by Healey and Picard as part of Healey's PhD work [3]. These authors find a predictive accuracy of 97.4% with LOOC for high stress and 94.7% for moderate stress based on data extracted from EMG, respiration, instantaneous heart rate (extracted from ECG), and GSR as well as additional features. The focus of this research was in generalized identification of stress states using signal fusion of multiple sensors, but it is not tailored for an individual for whom the stress is classified into three states. The unique aspect of our research is personalized individual stress analysis using ECG data alone. In addition, the data they used is not exactly the same so that results are not completely comparable. Yet our results to detect high stress are comparable to Healey and Picard [3].

Akbas calculated the differences between the 3 stress levels of averaged feature values extracted from instantaneous heart rate, EMG, hand GSR, foot GSR, instantaneous respiration rate, and average number of contractions per minute [9]. This author



**Table 5.** Sensitivity/Specificity and AUC of three classes with 10 selected attributes

Sens/Spec/AUC	Classification methods					
	Training set	Leave 1 out	2 Folds	10 Folds	75% Split	90% Split
Naïve Bayes	.731/ .878/.88	.567/ .783/.77	.582/ .773/ .753	.582/ .788/ .779	.588/ .752/ .835	.857/ .893/.976
BayesNet	.687/ .765/ .767	.672/ .773/ .581	.612/ .756/ .712	.657/ .768/ .713	.706/ .794/ .803	.857/ .893/.839
Logistic	.866/ .929/ .929	.716/ .834/ .830	.567/ .816/ .709	.657/ .818/ .834	.412/ .746/ .702	1/1/1
Multi. Perceptron	.94/.958/ .956	.716/ .827/ <b>.847</b>	.731/ .883/ <b>.815</b>	.806/ .899/ <b>.875</b>	.765/ .908/ <b>.932</b>	1/1/1
SMO	.776/ .883/ .853	.672/ .831/ .771	.627/.8/ .744	.701/ .848/ .795	.588/ .758/ .718	.857/ .893/.887
IB1	1/1/1	.642/ .779/ .711	.642/ .81/.726	.657/ .791/ .724	.647/ .832/ .739	.714/ .786/.75
IBK	1/1/1	.642/ .779/ .711	.642/ .81/.71	.657/ .791/ .765	.647/ .832/ .739	.714/ .786/.75
KStar	1/1/1	.537/ .753/ .689	.507/ .758/ .634	.552/ .752/ .693	.588/ .786/ .712	.857/ .893/1
ZeroR	.448/ .552/.5	.448/ .552/0	.448/ .552/ .492	.448/ .552/ .461	.294/ .706/.5	.429/ .571/.5
J48	.866/ .922/ .919	.731/ .851/ .711	.657/ .823/ .723	.672/ .822/ .731	.765/ .819/ .768	.857/ .893/.851
Random Forest	1/1/1	.701/ .815/ .827	.642/ .798/ .814	.687/ .815/ .842	.647/ .782/ .812	.857/893/ .893
Random Tree	1/1/1	.672/ .816/ .744	.612/ .812/ .712	.612/ .768/.69	.588/ .764/ .676	.857/ .893/.875

**Table 6.** SVMAttributeEval 10 selected features

1. avgDiffBeats	2. varDiffRR	3. varDiffQQ	4. varSS	5. avgBeats
6. varDiffSS	7. varDiffEntA10	8. varDiffEntA7	9. varEntA4	10. varEntA13

found these averages to be significantly different between the three levels of stress. However, no classification was performed.

Deng et al. extracted features from this data set based on principal component analysis (PCA) and determined that 5 features were best representative of this data set – foot GSR duration, hand GSR duration, hand GSR area, foot GSR area, and foot GSR frequency [12]. These results are not really comparable to ours since we are using only ECG. However, these authors are also classifying the signals using machine learning algorithms and report best average rate of 75.38% on all features with NaiveBayes and 78.46% accuracy on the 5 selected features with SVM, using LOOCV. These results are not exactly comparable to ours because the authors removed two additional samples, which can alter results on such a small data set. However, we concur that feature selection improves classification accuracy over not selecting features.

Deng et al. pursued their research by combining feature selection with signal selection, reducing the number of signals used to 2 [13] in their preceding study. They selected 5 features based on C4.5 and 2 sensors. With 10-fold CV (averaged 6 times), they obtained accuracy of 74.5% with SVM on all features and 85.46% with C4.5 on 5 features. However, they used only 65 data samples. The same authors published another paper applying combinatorial fusion to the same task [14] with comparable results. Since the data set used is not the same, their results are not comparable to ours.

Singh and Queyam also combined all sensors for the classification task using neural networks. They reported good results of over 80% on 6 out of 10 drives [15]. They reported on selecting features as being more correlated with driving conditions, and they found that mean heart rate and mean hand GSR were the most correlated [16]. However, they did not use entropy measurements so that their results are not clearly comparable to ours because they used multiple sensors.

Avki et al. reported also on correlations between features and stress level [17]. They found that the variance in the signals measurements is the most correlated to stress level, which confirms our results of the importance of entropy for classifying the signals. We also selected a number of variance features.

Some studies have focused on analyzing ECG signals alone to detect stress. Medina perform clustering and dimensionality reduction on raw signals to determine whether the learned clusters corresponded to stress levels [18]. This author reports good results, which our study corroborates. Her results are not directly comparable to ours since she is not performing supervised learning but only unsupervised learning. Moreover, her data set is different from ours. Another study, by Sun et al. focused on detecting mental stress based on ECG signals [19]. Also using a different data set, therefore not directly comparable to our work, the authors report best classification accuracy results of 92.4% using decision trees. This study therefore confirms the capability of tree models to best discriminate between the features during the classification task. Differences with our study include using a different data set, training and test sets at 50% split, and using galvanic skin response in addition to ECG, which they report as increasing the classification accuracy rate. Other studies on using sensors for stress detection are summarized in a review paper [20].

## 6 Conclusion

Using machine learning algorithms from the ECG signals alone, we could achieve up to 100% accuracy and 1 AUC, with Multilayer Perceptron, depending on the experimental setting, in detecting three classes of stress: low, medium and high. Thus the accuracy of detecting multiple stress levels based on individual variations in ECG extracted features is higher than that of previously published results detecting stress based on fiducial points alone. These results were obtained by adding multiscale entropy measurements in addition to the fiducial measurements performed in previous studies on the same data set. Future work will include adding the T-wave related features in our analysis since ECG studies have shown that QT is an important biomarker of cardiac abnormality [21] and adding other signals. Clearly the results presented here are limited by the small size of the data set (67 samples) so that studies on larger data sets need to be conducted. We also plan to work with a physician for future directions of this work and to explore additional classification and clustering algorithms, for example hierarchical methods.

Nevertheless, the results of the present study lead to the exciting possibility of monitoring and diagnosing individual stress levels and alert the users accordingly so that accidents committed due to high or prolonged stress can be prevented. The personalized signal classification analysis presented here can be extended to other situations in which people face stress thereby addressing fatigue in workers in a factory, failure in functioning of the elderly people, players in a field, soldiers in a war field, etc.

## References

1. Rubin, J., Eldardiry, H., Abreu, R., Ahern, S., Du, H., Pattekar, A., Bobrow, D.G.: Towards a mobile and wearable system for predicting panic attacks. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 529–533. ACM, September 2015
2. Cruz, L., Rubin, J., Abreu, R., Ahern, S., Eldardiry, H., Bobrow, D.G.: A wearable and mobile intervention delivery system for individuals with panic disorder. In: Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia, pp. 175–182. ACM, November 2015
3. Healey, J., Picard, R.W.: Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transp. Syst.* **6**(2), 156–166 (2005)
4. Rodrigues, J.G., Kaiseler, M., Aguiar, A., Silva Cunha, J.P., Barros, J.: A mobile sensing approach to stress detection and memory activation for public bus drivers. *IEEE Trans. Intell. Transp. Syst.* **16**(6), 3294–3303 (2015)
5. Keshan, N., Parimi, P., Bichindaritz, I.: Machine learning for stress detection from ECG signals in automobile drivers. In: Proceedings of International IEEE Conference on Big Data, pp. 2449–2457. IEEE (2015)
6. Price, D.: How to read and electrocardiogram (ECG). Part 1: basic principles of the ECG. The normal ECG. *South. Sudan Med. J.* **3**(2), 26–28 (2010)
7. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **3**, 230–236 (1985)

8. Goldberger, A.I., Amara, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* **101**(23), e215–e220 (2010). doi:10.1161/01.CIR.101.23.e215. *Circulation Electronic Pages*: <http://circ.ahajournals.org/cgi/content/full/101/23/e215>. 13 June 2000. PMID: 10851218
9. Akbas, A.: Evaluation of the physiological data indicating the dynamic stress level of drivers. *Sci. Res. Essays* **6**(2), 430–439 (2011)
10. Hall, M., Frank, E., Holmes, G., Pfahringer, G., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. *SIGKDD Explor.* **11**(1), 10–18 (2009)
11. Costa, M., Goldberger, A.L., Peng, C.-K.: Multiscale entropy analysis of biological signals. *Phys. Rev.* **71**, 021906-1–021906-18 (2005)
12. Deng, Y., Wu, Z., Chu, C.-H., et al.: Evaluating feature selection for stress identification. In: *IEEE Proceedings of the 13th International Conference on Information Reuse and Integration (IRI)*, pp. 584–591. IEEE (2012)
13. Deng, Y., Wu, Z., Chu, C.-H., et al.: An investigation of decision analytic methodologies for stress identification. *Int. J. Smart Sens. Intell. Syst.* (2012). ISSN: 1178-5608
14. Deng, Y., Wu, Z., Chu, C.-H., et al.: Sensor feature selection and combination for stress identification using combinatorial fusion. *Int. J. Adv. Rob. Syst.* **10**, 306–313 (2013)
15. Singh, M., Queyam, A.B.: A novel method of stress detection using physiological measurements of automobile drivers. *Int. J. Electron. Eng.* **5**(2), 13–20 (2013)
16. Singh, M., Queyam, A.B.: Correlation between physiological parameters of automobile drivers and traffic condition. *Int. J. Electron. Eng.* **5**(2), 6–12 (2013)
17. Avci, C., Akbas, A., Yüksel, Y.: Evaluation of Statistical Metrics by Using Physiological Data to Identify the Stress Level of Drivers
18. Medina, L.: Identification of stress states from ECG signals using unsupervised learning methods. Dissertation, Master's thesis, Universidade Técnica de Lisboa, Instituto Superior Técnico (2009)
19. Sun, F.T., Kuo, C., Cheng, H.T., Buthpitiya, S., Collins, P., Griss, M.: Activity-aware mental stress detection using physiological sensors. In: *Mobile Computing, Applications, and Services*, pp. 211–230. Springer, Heidelberg (2010)
20. Begum, S.: Intelligent driver monitoring systems based on physiological sensor signals: a review. In: *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems-(ITSC)*, pp. 282–289. IEEE (2013)
21. Zhou, Y., Sedransk, N.: Functional data analytic approach of modeling ECG T-wave shape to measure cardiovascular behavior. *Ann. Appl. Stat.* **3**, 1382–1402 (2009)