Machine Learning Ranks ECG as an Optimal Wearable Biosignal for Assessing Driving Stress

MOHAMED ELGENDI[1,2,3], **(Senior Member, IEEE),**
AND CARLO MENON[1], **(Senior Member, IEEE)**

[1]Menrva Research Group, Schools of Mechatronic Systems Engineering and Engineering Science, Simon Fraser University, Surrey, BC V3T 0A3, Canada
[2]Faculty of Medicine, The University of British Columbia, Vancouver, BC V6T 1Z3, Canada
[3]School of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Corresponding author: Mohamed Elgendi (moe.elgendi@gmail.com)

ABSTRACT The demand for wearable devices that can detect anxiety and stress when driving is increasing. Recent studies have attempted to use multiple biosignals to detect driving stress. However, collecting multiple biosignals can be complex and is associated with numerous challenges. Determining the optimal biosignal for assessing driving stress can save lives. To the best of our knowledge, no study has investigated both longitudinal and transitional stress assessment using supervised and unsupervised ML techniques. Thus, this study hypothesizes that the optimal signal for assessing driving stress will consistently detect stress using supervised and unsupervised machine learning (ML) techniques. Two different approaches were used to assess driving stress: longitudinal (a combined repeated measurement of the same biosignals over three driving states) and transitional (switching from state to state such as city to highway driving). The longitudinal analysis did not involve a feature extraction phase while the transitional analysis involved a feature extraction phase. The longitudinal analysis consists of a novel interaction ensemble (INTENSE) that aggregates three unsupervised ML approaches: interaction principal component analysis, connectivity-based clustering, and K-means clustering. INTENSE was developed to uncover new knowledge by revealing the strongest correlation between the biosignal and driving stress marker. These three MLs each have their well-known and distinctive geometrical basis. Thus, the aggregation of their result would provide a more robust examination of the simultaneous non-causal associations between six biosignals: electrocardiogram (ECG), electromyogram, hand galvanic skin resistance, foot galvanic skin resistance, heart rate, respiration, and the driving stress marker. INTENSE indicates that ECG is highly correlated with the driving stress marker. The supervised ML algorithms confirmed that ECG is the most informative biosignal for detecting driving stress, with an overall accuracy of 75.02%.

INDEX TERMS Digital health, anxiety monitoring, mental well-being, biosignals, mobile health, wearable devices.

I. INTRODUCTION

The American Psychological Association has reported that people are currently living with extremely high levels of driving stress, and this is expected to increase in the coming years [1]. Recently, researchers have found that stress plays a major role in the development and progression of cardiovascular diseases [2]. Therefore, the early detection and continuous assessment of stress can improve health and well-being, helping to prevent severe consequences that can lead to hospitalization and even death.

Recent advances in wearable devices, biosignal processing, machine learning (ML), and app development could contribute to providing objective feedback on driving stress. Wearable devices have already been used in several health interventions, achieving promising results [3].

The most used and best understood, biosignal in modern medicine is the electrocardiogram (ECG). Thus, ECG-based wearable devices have a great potential to succeed. Note that most of the available ECG wearable devices on the market are

The associate editor coordinating the review of this manuscript and approving it for publication was Ying Xu.

not approved by the Food and Drug Administration. Several studies have investigated ECG signals that detect different reactions to anxiety, reporting contradictory results [4]. However, limited research has been conducted to compare the efficacy of using ECGs with other biosignals in detecting stress. In 2018, El Haouij [5] developed an algorithm to detect stress using the ECG, electromyogram (EMG), Palmar galvanic skin resistance (Palmar GSR), foot galvanic skin resistance (Plantar GSR), heart rate (HR), and respiration (RESP). She reported a 69% accuracy for detecting stress. It is worth noting that El Haouij has investigated the importance of each biosignal; however, not all biosignals were included in her investigation, such as the ECG, which was removed. Recently, Smets et al. [6] developed an algorithm that used an ECG, skin conductance, and skin temperature to detect driving stress, eventually reaching 43% accuracy. However, the study was not focused on questioning the importance of each biosignal or ranking the biosignal based on its driving stress detection accuracy. A recent study by Rizwan et al. [7] used only ECG signals to assess driving stress; however, their analysis was carried out independently without reporting a relative comparison with other biosignals.

To the best of our knowledge, limited efforts have been made to determine the effectiveness of the different biosignals that can be collected by wearable devices, with a focus on the ECG signal as a potential single biosignal for detecting driving stress. In the current study, we will uncover hidden relationships among all biosignals and investigate the importance of each biosignal for assessing driving stress.

II. MATERIALS AND METHODS
A. DRIVING STRESS DATASET
The Stress Recognition in Automobile Drivers (SRAD) database [8] is useful because it contains a collection of multiparameter biosignals collected simultaneously from healthy volunteers. The biosignals were collected using wearable devices while the volunteers were driving on a prescribed route including city streets and highways in and around Boston, Massachusetts, USA. Six different biosignals were collected in the SRAD database: ECG, EMG, Plantar GSR, Palmar GSR, HR, and RESP, all during real time; In addition, the driving stress marker time series was collected as a way to reflect the driver's level of stress. The database is publicly available and can be accessed via a website [9]. All signals were recorded over seven driving periods (or segments) that are as follows: first rest (R1), first drive in the city (C1), first drive on the highway (HW1), second drive in the city (C2), second drive on the highway (HW2), third drive in the city (C3), and final rest (R2). The reason we chose the SARD database is that to the best of our knowledge, there are no other publicly available databases that contain biosignals that were simultaneously collected to assess driving stress. Moreover, scientific evidence suggests that work-related stress, life stress, and driving environment stress impact on driving outcomes [10]. Thus, analysis of the SRAD may be valuable

in identifying the optimal biosignal for driving stress assessments in a practical multimodal stress scenario.

B. EXCLUSION CRITERIA
The SRAD database contains biosignals collected from 17 drivers. The experiment was quite complex, and the duration of the driving varied between 50 and 90 minutes. All drivers started with approximately 15 minutes of rest, after which the driver drove the car out of the garage toward a congested city street. Full consistency in collecting data and labeling driving stress was difficult to achieve; unfortunately, not all signals were provided for all the drivers. Moreover, in a few recordings, the marker signals were not clear. Thus, nine recordings were excluded, as shown in Table 1. Eight recordings with complete signals and clear driving stress markers were finally included in the analysis.

TABLE 1. Excluded recordings from the analysis.

Excluded Record	Reason
$drive01$	Marker is missing
$drive02$	Palmar GSR and EMG are missing
$drive03$	HR & EMG and Marker are missing
$drive04$	EMG is missing
$drive09$	Marker is not clear
$drive13$	Palmar GSR is missing
$drive14$	HR is missing
$drive16$	Marker is not clear
$drive17$	Marker is not clear

C. METHODS
1) HYPOTHESIS
It is assumed that an individual's stress level increases when driving in the city and decreases when not driving (at rest) [11]. Two approaches for assessing driving stress must be investigated: longitudinal stress (repeated measurement of the same biosignals combined over three driving states) and transitional stress (switching from state to state, such as city to highway driving). Thus, our hypotheses are as follows:

1) Based on the longitudinal stress analysis, the biosignal most highly correlated with the marker signal (regardless of whether the driver was in the city or on the highway) will be the most valuable and informative in an assessment of driving stress.
2) Based on the transitional stress analysis, the biosignal that provides the highest driving stress classification accuracy will be best used for a driving stress assessment.
3) Based on the multicomparison stress analysis, the biosignal that most separates the three driving states should be used for assessing driving stress.
4) The biosignal that satisfies the previous three analyses will be considered the optimal biosignal for assessing driving stress.

2) LONGITUDINAL STRESS ANALYSIS
The longitudinal stress analysis is the combined repeated measurement of the same biosignals over three driving states.

In this analysis, we examined all collected biosignals with and without feature extraction. Note that the biosignals are filtered (bandpass 0.5–7 Hz Butterworth filter). We propose a novel interaction ensemble called INTENSE that aims to combine multiple clustering models to produce robust results. Note that INTENSE will be used for all collected biosignals as is, without feature extraction. Many clustering models have been reported in literature. Each method has a different set of rules for defining "similarities" among biosignals. We selected clustering methods that are commonly used for simplicity and effectiveness. INTENSE aggregates results from the following three clustering models, which are based on three different geometrical perspectives:

1) The interaction principal component analysis, a multi-interaction analysis, produces multiple biosignals that interact with each other. This analysis represents a correlation-based eigenvector-based multivariate analysis.
2) Connectivity-based clustering, a bi-interaction analysis, groups every two biosignals based on the similarity measure (pairwise distance) used.
3) K-means clustering, which is a centroid-based interaction analysis, groups all biosignals based on the similarity measure (distances between centroids) used.
4) The ensemble method aggregates the results from the three clustering methods to produce a combined recommendation about the interactivity among all biosignals in the dataset.

INTENSE was developed because each clustering method has potential shortcomings around clustering a certain dataset depending on the initial value used, thresholds, similarity measure, etc. The use of only one clustering algorithm may provide insights into the interactions between biosignals, based on the algorithm's geometrical concept. However, if we combine the results from multiple clustering methods, each one affords a unique geometrical perspective. This will allow us to obtain solid findings and to agree upon different clustering methods using a majority voting rule.

3) INTERACTION PRINCIPAL COMPONENT ANALYSIS

The interaction principal component analysis (IPCA) proposed in Ref [13] was used in this study. IPCA is a correlation-based unsupervised machine learning (ML) technique. It finds linearly uncorrelated attributes within a set of observations of possibly correlated attributes (in our case, biosignals). It starts with a decorrelation process that does not requires any prior information or settings about the processed biosignals, followed by Pearson's correlation process. IPCA is a valuable technique for automatically revealing hidden interactions between biosignals, without training or labeling. IPCA analysis provides a level of artificial intelligence for uncovering new behaviors between examined biosignals. The IPCA algorithm involves the following steps:

1) All recordings were combined into one matrix, X, with the dimensions of (764, 483 samples × 7 biosignals).

2) The Min-Max normalization was applied, as follows $Z = (X - X_{\min})/(X_{\max} - X_{\min})$, this was done to ensure that Z had a zero mean.
3) The covariance matrix was obtained $C = \frac{1}{7}\sum_{f=1}^{7} ZZ^T$, where f is the number of biosignals.
4) The eigen decomposition of C was performed and the e_f eigenvalues and their corresponding eigenvectors, v_f, were computed to satisfy the equation $Cv = ev$.
5) The eigenvalues were sorted in descending order, $e_1 \geq e_2 \geq \ldots \geq e_7$, along with their corresponding eigenvectors. Note that the eigenvector with the largest eigenvalue was called the first principal component (PC1) while the last PC associated with the lowest eigenvalue was called PC7.
6) We calculate Pearson's correlation between biosignals and between the biosignals and PCs.
7) We select PCs that show more than two strong correlations ($|r > 0.5|$) with variables (biosignals), and then return these variables (biosignals).

Seven segments (or periods) were extracted from the seven biosignals: a driving stress marker, ECG, EMG, Plantar GSR, Palmar GSR, HR, and RESP. To compare the different driving segments (e.g., driving in the city vs. driving on the highway), the Wilcoxon–Mann–Whitney test (p_w) was used for two independent groups. A p value of < 0.05 was considered significant. Pearson's correlation coefficient was used to calculate the correlation between the biosignals and PCs. We used Matlab 2018b software and Python 3.6.5 software to analyze the data.

4) CONNECTIVITY-BASED CLUSTERING

The connectivity-based clustering (CBC) is an algorithm that connects "biosignals" to form "groups" based on their distance. It is a different unsupervised ML approach that was utilized to quantify and visualize the dissimilarities among the biosignals. The Euclidean distance was used as a metric to provide hierarchical clustering, which is also called a Dendrogram.

5) K-MEANS CLUSTERING

K-means clustering (KMC) is considered to be the most widely known clustering algorithm as its concept is simple but effective. This algorithm divides the number of factors F into C disjoint clusters. The means of the clusters are referred to as "centroids." The KMC algorithm aims to choose centroids that minimize the distance between each group of factors and its centroid. The minimization process is called "inertia" or within-cluster sum-of-squares, and it is defined as follows: $||f_i - c_j||^2$, where f_i refers to all values in factor i, and c_j refers to cluster j. Note that applying PCA prior to KMC clustering is highly recommended to speed up the process and avoid computational problems.

6) MULTICOMPARISON TEST

Here, the multiple statistical comparison test was based on the mean square error achieved by the Kruskal-Wallis test.

The multicomparison displays the mean estimates and standard errors with the corresponding state. Each driving state mean is represented by a symbol, and the interval is represented by a line extending out from the symbol. Two driving state means are significantly different if their intervals are disjointed; they are not significantly different if their intervals overlap. Note that the biosignals were not used as is, which is in contrast to the INTENSE analysis. A feature extraction was applied first, as follows:

The Skewness statistic measure was used as a feature to compare all the biosignals because a previous study [14] found that Skewness is associated with morphology changes in time-series biosignals. Note that the study [14] compared different features and ranked Skewness as the optimal feature. Skewness is a measure of the symmetry (or the lack thereof) of a probability distribution, which is defined as follows:

$$S = 1/N \sum_{i=1}^{N} [x_i - \widehat{\mu}_x/\sigma]^3, \qquad (1)$$

where $\widehat{\mu}_x$ and σ are the empirical estimate of the mean and standard deviation of x_i, respectively, and N is the number of samples in the biosignal.

7) TRANSITIONAL STRESS ANALYSIS

The transitional stress analysis is defined as switching from state to state, specifically for the following five transitional states: (R1 vs. C1), (H1 vs. C1), (R1 vs. C2), (H2 vs. C2), and (R2 vs. C3). The transitional stress analysis relies on supervised ML algorithms. Using a leave-on-out cross-validation, 17 different supervised classification techniques were tested: KNN weighted, Tree Fine, SVM Cubic, Logistic Regression, SVM Quadratic, Quadratic Discriminant, and SVM Fine Gaussian. Skewness was applied as a feature to capture driving stress.

An algorithm is needed to detect state switching in order to automatically extract biosignals based on the seven driving states. The two event-related moving averages (TERMA) algorithm [12] was used in our study to detect the main spikes in the marker signal.

8) TERMA METHOD

TERMA is computationally efficient and robust to noise. Therefore, the TERMA was used to segment each biosignal into seven groups: R1, C1, HW1, C2, HW2, C3, and R2.

$$MA_i[n] = \frac{1}{W_i} \sum_{n=1}^{N} \sum_{m=n-(W_i-1)/2}^{M=n+(W_i-1)/2} y[m], \qquad (2)$$

$$THR_1 = MA_2[n] + \alpha, \quad THR_2 = W_1. \qquad (3)$$

$$BLOCKS[n] = MA_1[n] > THR_1, \qquad (4)$$

$$Peaks = max(BLOCKS[n] > THR_2), \qquad (5)$$

The TERMA-based marker detection function has five inputs (starting frequency [F_1], end frequency [F_2], first moving average [MA_1 with a window size of W_1], second moving average [MA_2 with a window size of W_2], and rejection

threshold [β]). First, the driving stress marker signal passes through a third-order Butterworth bandpass filter F_1–F_2. The resulting signal is squared, and two moving averages (MA_1 and MA_2) are applied, followed by a threshold (β to reject noise), this process generates the so-called "blocks of interest". By searching for the maximum absolute value within each block of interest, the marker location can be determined. Here, we found that the optimal setting for detecting markers was the following: $F_1 = 0.5$ Hz, $F_2 = 7$ Hz, $W_1 = 500$ samples, $W_2 = 1000$ samples, and $\beta = 5$.

9) DATA AVAILABILITY

The data that support the findings of the current study are publicly available online at https://www.physionet.org/pn3/drivedb/

III. RESULTS

The overall visual representation of biosignals for *drive*07 is shown in Fig. 1. It is very difficult to differentiate among the biosignals collected during the city driving, highway driving, and at rest by visual inspection alone. Even if we can identify some patterns or significant changes between the different segments or periods in one recording (or one biosignal), we may not be able to generalize the findings.

There was a need to automatically segment all biosignals based on the driving stress marker signal to statistically compare all segments. Thus, the TERMA method was used to segment the biosignals into seven periods: R1, C1, HW1, C2, HW2, C3, and R2. The TERMA method was applied to detect the switching states within the driving stress marker, as shown in Fig. 2. Based on visual inspection, the detection accuracy of the seven periods achieved on all used records was 100%. This step was important to demarcate the driving activity in all biosignals for comparative purposes.

Fig. 3 shows the significance of the seven PCs extracted from the SRAD database, after applying a principal component analysis (PCA) to all biosignals (driving stress marker, ECG, EMG, Plantar GSR, Palmar GSR, HR, and RESP). PC1 explains the most variance (40%), reflecting its relevance and importance. PC2 explains nearly 23% of the variance, while PC3 explains a little more than 17% of the variance. As expected, upon visual inspection, PC1 is the most volatile PC2 is the second most volatile, and PC7 is the least volatile. In fact, PC7 showed no relevance.

Fig. 4 shows the heat map of the IPCA longitudinal stress analysis, which is a correlation matrix among all PCs and biosignals. The diagonal entries are all equal to 1. As can be seen, the heat map consists of four 7×7 blocks. The top right block is the correlation matrix for the PCs. We know that the PCs are orthogonal. As expected, this block contains zeros (numerically negligible values) in its off-diagonal entries, demonstrating that the PCs are mutually orthogonal (and therefore uncorrelated). The bottom left block is the heat map showing no correlation between all biosignals. Thus, the bottom left block is not informative.
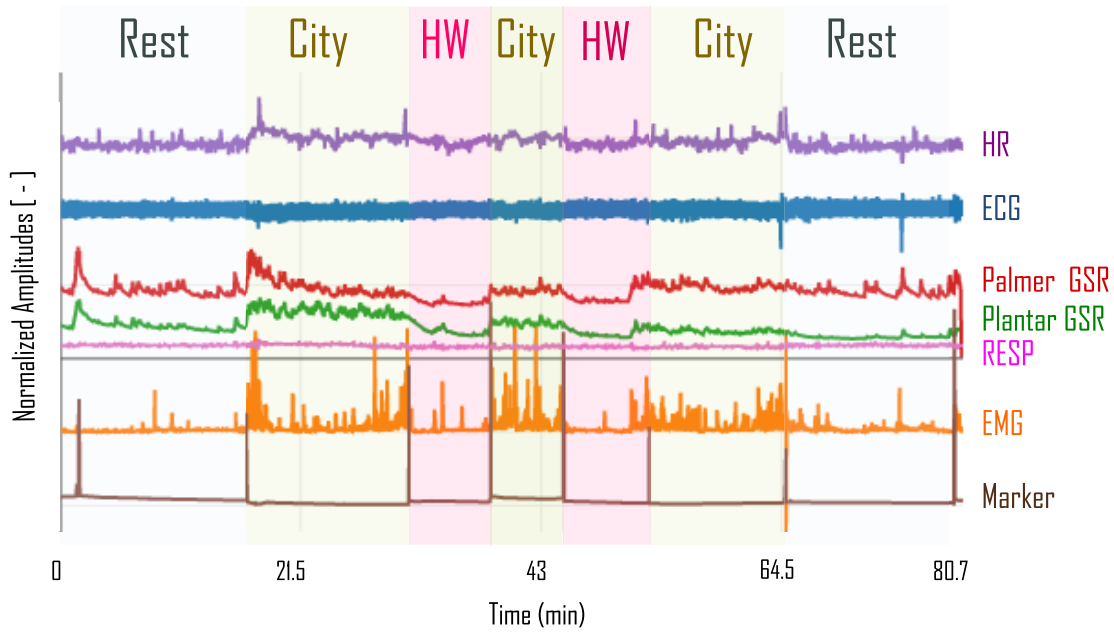
FIGURE 1. Visualizing all biosignals collected from volunteer #7 (*drive*07) during three driving states: city, highway, and at rest.
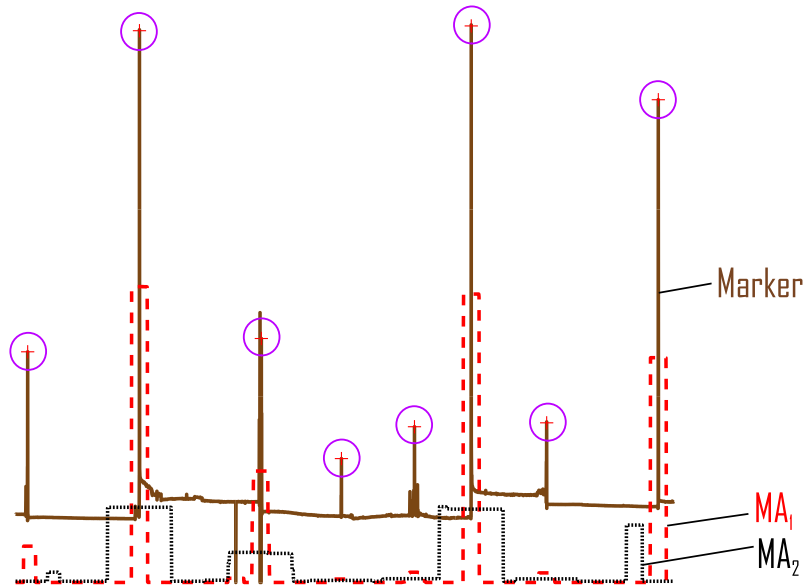

FIGURE 2. Detection of state switching of the driving stress marker using the TERMA (two event-related moving averages) method [12]. This is an important step for the longitudinal and transitional stress analysis.

As shown in Fig. 4, the bottom right 7×7 block contains interesting results. The 7×7 block, which is surrounded by dashed black lines, reflects the correlation between each biosignal and each PC. The first column, in the dashed block, shows a significant correlation between PC1 and the marker, ECG, and RESP. PC1 shows that ECG and RESP are correlated with the driving stress marker; however, both features are moving in the same direction of the driving stress

marker. Given that PC1 is the PC with the highest eigenvalue, there is a hidden correlation among the ECG, RESP, and the driving stress marker. The second column, which is in the dashed block, shows that the Palmar GSR and Plantar GSR are moving in the same direction, and both are strongly correlated with PC2. This suggests that there is a hidden correlation between the Palmar GSR and Plantar GSR. The third column, which is in the dashed box, shows that the
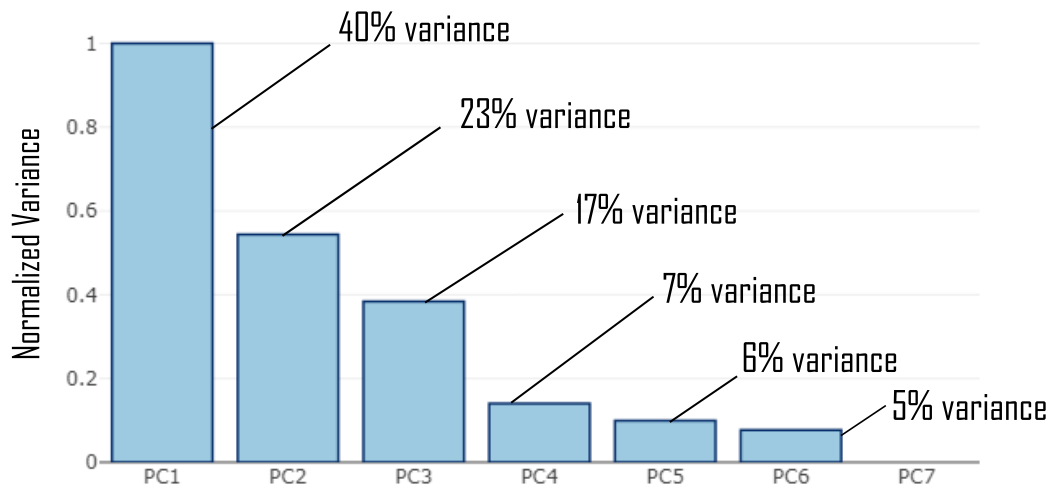
FIGURE 3. Principal component analysis (PCA) of all biosignals used from the Stress Recognition in Automobile Drivers (SRAD) database. PC1 has the highest variance, while PC7 has the lowest variance.

driving stress marker and ECG move in the same direction, and both are strongly correlated with PC3, suggesting that a hidden correlation exists between the driving stress marker and ECG. PC4 is only correlated with HR while PC5 is correlated with EMG and HR. PC6 is correlated with Plantar GSR while PC7 was unable to capture any correlation between biosignals. Interestingly, this analysis highlights the correlation between the driving stress marker with ECG, as captured by PC1 and PC3.

Here, we investigated the hierarchical clustering of biosignals and analyzed the Dendrogram generated as a result of the clustering process. This method builds the hierarchy from the independent biosignals by progressively merging the biosignals to generate clusters. In this case, we have six biosignals. The first step is to determine which two biosignals are close enough to each other to be merged into a cluster.

The biosignals were grouped into clusters using the longitudinal CBC analysis, based on their similarity, as shown in Fig. 5. Visually inspecting Fig. 5, the changes in the driving stress marker are similar to those in ECG, EMG, Palmar GSR, and HR. Interestingly, the first biosignal close to the driving stress marker is ECG. Moreover, the Palmar GSR and Plantar GSR are forming an independent cluster far from the driving stress marker. These results are similar to the findings from the PCA shown in Fig. 4, hence confirming the correlation between Palmar GSR and Plantar GSR. Interestingly, the Dendrogram ranked the biosignals based on how close they were to the driving stress marker in the following order: ECG, RESP, EMG, Plantar GSR, Palmar GSR, and HR.

KMC serves as our third geometrical lens and provides a different clustering perspective. The KMC result, as shown in Fig. 6, is similar to and confirms the findings obtained using the CBC, as shown in Fig. 5. Interestingly, KMC grouped the stress marker with ECG and EMG, suggesting that the marker is correlated with ECG and EMG. Therefore, INTENSE indicates that ECG is consistently correlated with

the stress marker, suggesting that ECG is an optimal biosignal for assessing driving stress.

It is necessary to examine the pairwise comparison results using a multiple comparison test for each biosignal in terms of city driving, highway driving and at rest. The longitudinal multiple comparison can generate a graph of the mean estimates and standard errors. Fig. 7 shows the multiple comparison of the means of the extracted feature (Skeweness) for each biosignal. The mean of each status (e.g., at rest) is represented by a symbol, and the interval is represented by a line extending from the symbol.

As can be seen in Fig. 7 the mean of at rest is highlighted and the comparison interval is in blue. As the comparison intervals for driving in the city and on the highway do not intersect with the intervals for the at rest mean, they are highlighted in red, except for in Fig. 7(HR). The intersection among the intervals indicates no significant difference, while the lack of intersection indicates that both means are different than the at rest mean. In other words, if the means of driving in the city and on the highway are significantly different and if their intervals are disjointed, they are not significantly different if their intervals overlap.

Fig. 7 shows that HR is the worst signal for differentiating among the three driving statuses. The rest of the biosignals can differentiate being at rest and either driving in the city or on the highway. However, ECG is the only biosignal with no overlap between driving in the city and on the highway. This result suggests that an ECG is the optimal biosignal for capturing stress in different driving statuses.

Fig. 7 magnifies the separability among the biosignals, suggesting a need to examine the classification ability of each biosignal. Table 2 summarizes the transition stress analysis, showing the classification accuracy between the transitional states: (R1 vs. C1), (H1 vs. C1), (R1 vs. C2), (H2 vs. C2), and (R2 vs. C3). Interestingly, the ECG and RESP scored the highest accuracy of 81.3% in detecting driving stress in the
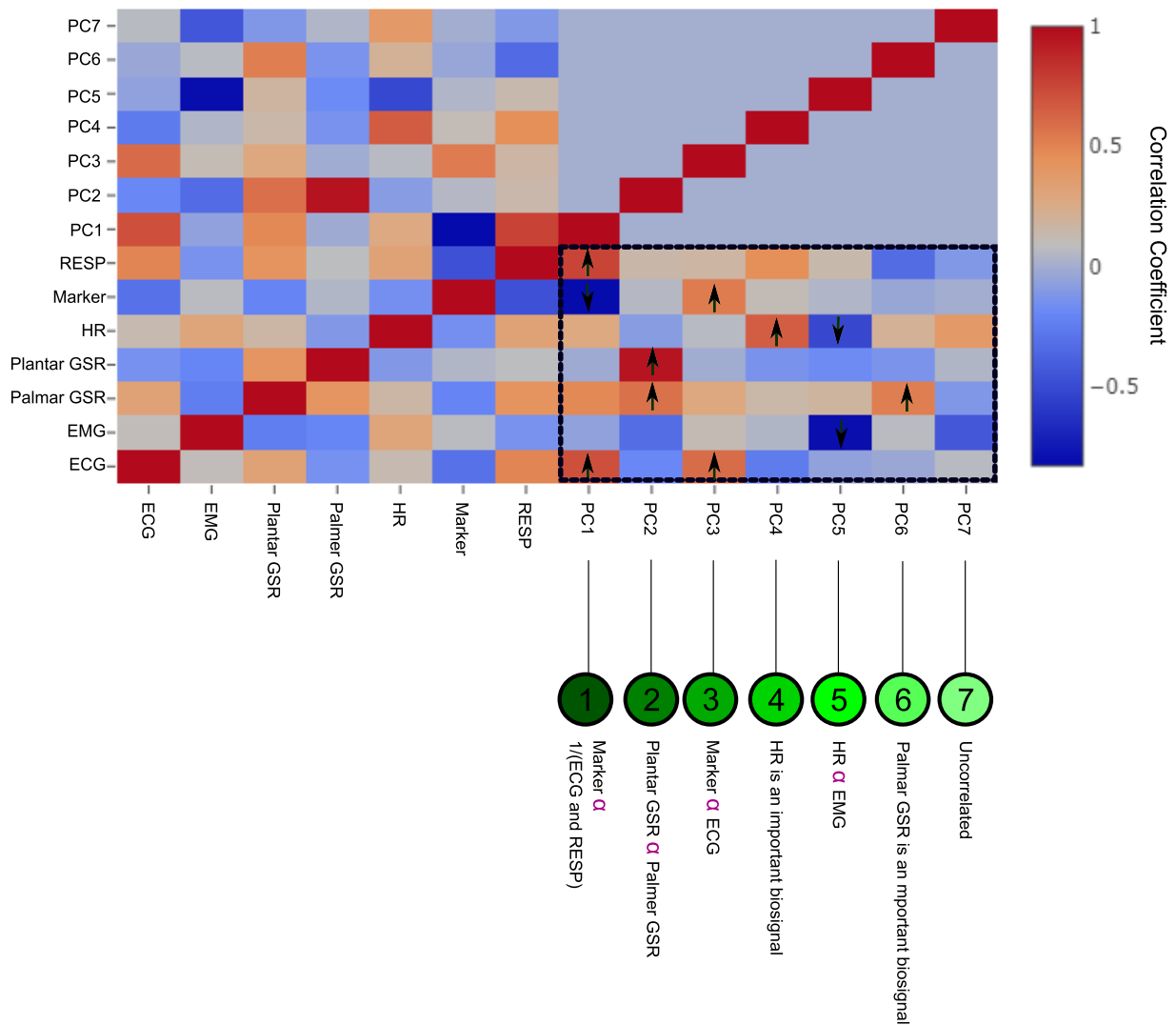
FIGURE 4. Longitudinal stress analysis using the interaction principal component analysis (IPCA) for all biosignals obtained from the Stress Recognition in Automobile Drivers (SRAD) database. Here, α refers to the hidden correlation. The black arrow indicates an absolute value of the correlation coefficient $>= 0.5$ is achieved.

test (R1 vs. C1) using the KNN-Weighted classifier. Plantar GSR scored the highest accuracy in detecting driving stress in the test (R2 vs. C3) using the Quadratic classifier, while RESP was the most accurate (R1 vs. C2) using the Tree-Fine classifier.

It is clear that ECGs were able to capture overall driving stress in all possible driving status tests, with an overall accuracy of 75.02%. The second highest accuracy was achieved by Palmar GSR with an overall score of 72.52%, while the lowest accuracy was scored by HR, with an overall accuracy of 63.78%.

IV. DISCUSSION
A. DRIVING STRESS MARKER
The driving stress marker is a continuous time-series signal that divides various driving zones due to the assumption that driving in city evokes high stress levels, driving on

the highway evokes medium stress levels, and that being in a state of rest evokes the lowest stress levels. The driving zone marker in the current study is referred to as the driving stress marker. The driving stress signal is typically used as the gold standard for assessing different driving stress level changes over time for each recording. If the quality of the driving stress marker is high, there will be eight clear outstanding spikes that determine seven driving zones (R1, C1, H1, C2, H2, C3, and R2), as shown in Fig. 2.

B. ECG
An ECG is a time-series signal that reflects the electrical activity of the heart [15]. Usually, three electrodes are placed on the chest to record ECG signals. In the original study, the ECG signals were captured using a modified lead II configuration to minimize motion artifacts and to attain a
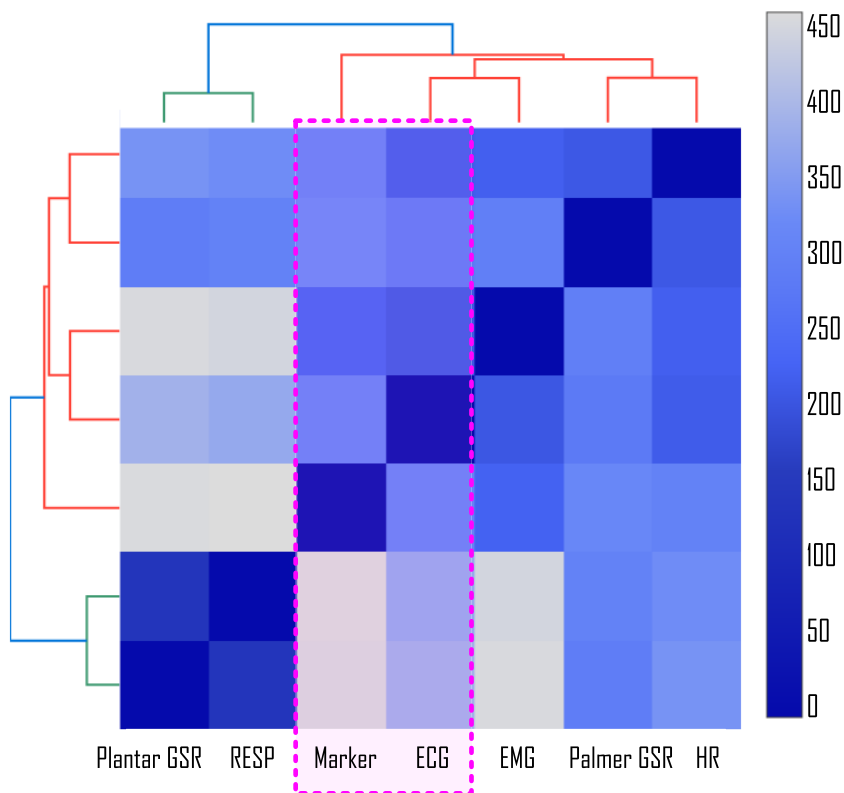
FIGURE 5. Longitudinal stress analysis using connectivity-based clustering for all biosignals obtained from the Stress Recognition in Automobile Drivers (SRAD) database.
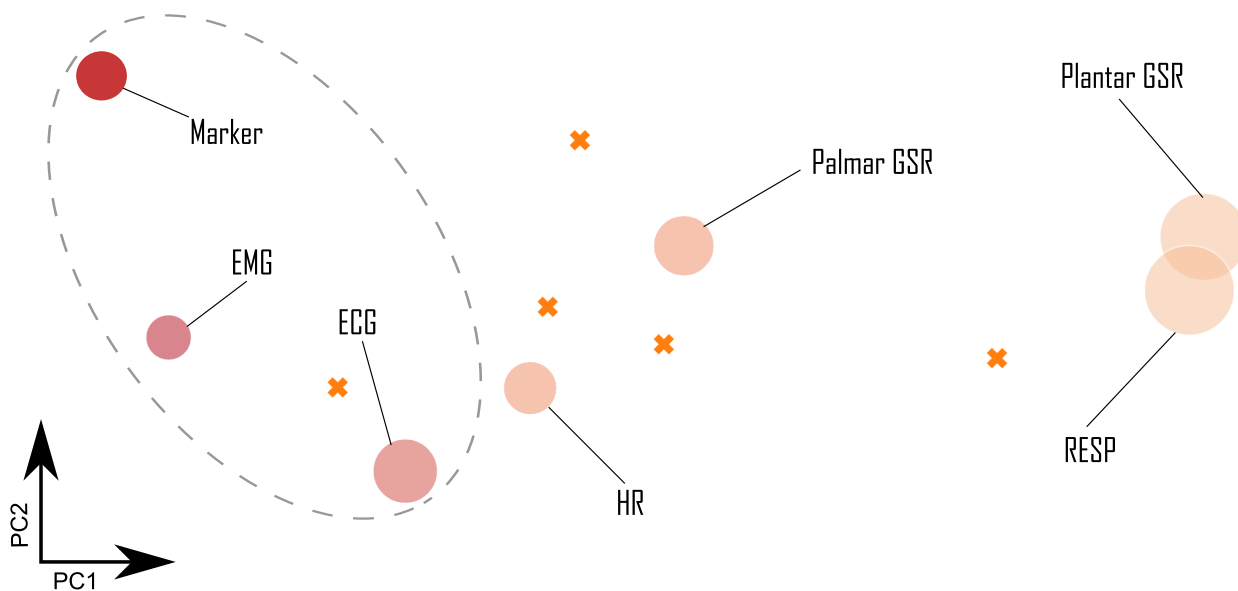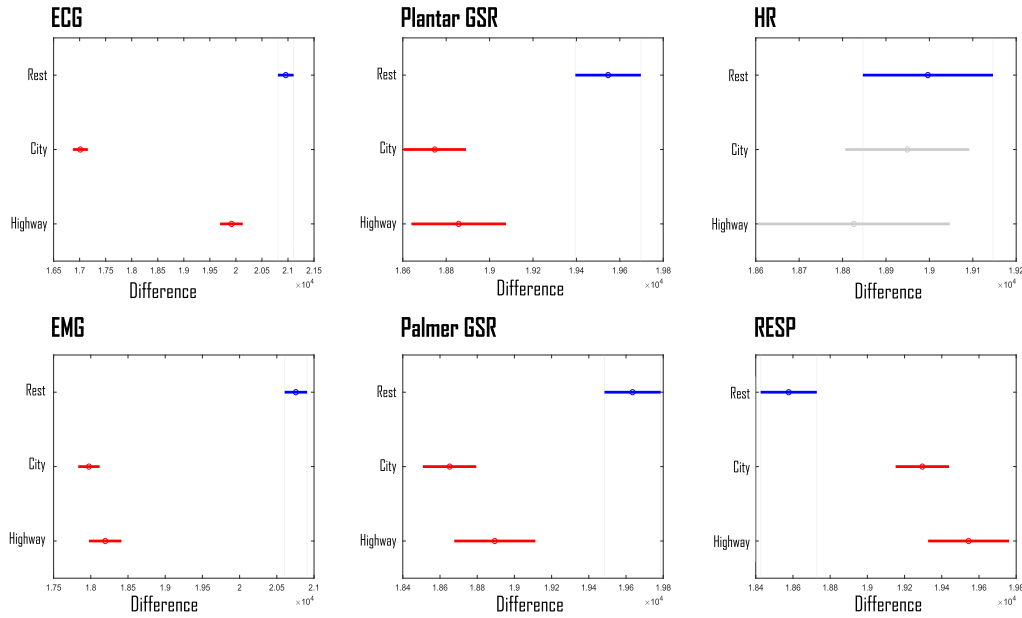

FIGURE 6. Longitudinal stress analysis using K-means clustering (KMC) for all biosignals obtained from the Stress Recognition in Automobile Drivers (SRAD) database. Here, 'x' represents the centroid of a cluster while PC stands for principal component.

good record of the R peak. Previous studies investigated the correlation of different ECG features with stress by using the same dataset, such as the R peak amplitude [16], QT interval [17], and QRS complex [18]. These studies can

be considered pilot studies without a clear discussion of how the features were extracted. In addition, no cross validation was used, and no relative analysis to other biosignals was reported.

FIGURE 7. Longitudinal stress analysis of the multicomparison test for all biosignals (Skewness with one second window was used as a feature). Two driving periods are significantly different if their intervals are disjointed; the two driving periods are not significantly different if their intervals overlap with a gray color. One can observe that ECG significantly differentiates between three driving states. Note that EMG, RESP, Plantar GSR, and Palmar GSR overlap with driving in the city and highway while HR has a complete overlap with all driving states.

TABLE 2. Transitional stress analysis using supervised machine learning (ML) algorithms. It compares the classification accuracy among driving states using biosignals. Note that Skewness with one second window was used as a feature. Leave-one-out cross validation was used.

	R1 vs. C1 (%)	H1 vs. C1 (%)	R1 vs. C2 (%)	H2 vs. C2 (%)	R2 vs. C3 (%)	Overall Accuracy (%)
ECG	81.3 KNN weighted	75 Tree Fine	75 Tree Fine	75 Tree Fine	68.8 KNN weighted	75.02
EMG	68.8 Tree Fine	62.5 Logistic Regression	62.5 Cubic SVM	56.3 Logistic Regression	75 Medium Gaussian SVM	65.02
Plantar GSR	75 SVM Fine Gaussian	50 Cubic SVM	68.8 Linear SVM	56.3 Linear SVM	81.3 Quadratic Discriminant	66.28
Palmar GSR	68.8 Cubic SVM	87.5 KNN Fine	87.5 KNN weighted	56.3 Cubic SVM	62.5 Quadratic Discriminant	72.52
HR	50 Tree Fine	62.5 Quadratic SVM	68.8 Quadratic Discriminant	68.8 Quadratic Discriminant	68.8 Quadratic SVM	63.78
RESP	81.3 Tree Fine	50 KNN Cosine	81.3 Tree Fine	50 KNN Cosine	62.5 Quadratic Discriminant	65.02

In our study, the ECG was analyzed in relation to other biosignals, and all biosignals were included in the records. The longitudinal IPCA analysis showed that the ECG is correlated with the driving stress marker and RESP. In addition, the Dendrogram showed that ECG was ranked as closest to the driving stress marker. The multiple comparison of means reflected a significant difference for all driving status tests when using ECG, without any overlap between the three driving statuses, as shown in Fig. 7. Interestingly, ECG was ranked first for detecting driving stress based on the overall accuracy (75.02%), as shown in Table 2. To our knowledge, this finding has *not* been reported in the literature; in fact, some features extracted from ECG signals have led to contradictory results [4] in detecting anxiety. It seems that the morphology of the ECG waveforms contains valuable information and is correlated with driving stress.

C. EMG

An EMG noninvasively measures muscle action potentials during contraction. In the literature, EMGs have been investigated for analyzing six emotional states [19], identifying intensive valence and arousal affective states [20] and detecting driving stress [8], [21]. Our study examined the effectiveness of EMG relative to other biosignals. The IPCA analysis showed that EMGs are not correlated with the driving stress marker; however, EMG was correlated with HR. The Dendrogram showed the EMG was ranked third in closeness to the driving stress marker. The multiple comparison of means reflected a significant difference for all driving status tests using EMGs. However, there was an overlap between the city and highway driving, as shown in Fig. 7. After applying different supervised classifiers, EMG was ranked fourth for detecting driving stress based on the overall accuracy

(65.02%), as shown in Table 2. This finding has been confirmed previously in the literature; in fact, a previous study [5] showed that EMGs were the least informative biosignal for detecting driving stress.

D. PLANTAR GSR

Plantar GSR is the galvanic skin response or skin conductance measured by the foot. A GSR measures autonomic nervous system activity. Typically, two electrodes are placed on the foot skin to capture changes in the skin's glands when producing ionic sweat [22]. It worth noting that the Plantar GSR was not considered in the original study [8], [11]. However, a recent study [5] showed that the Plantar GSR was the most biosignal correlated with driving stress. We considered the Plantar GSR in our study and examined its importance. The IPCA analysis showed that the Plantar GSR and Palmar GSR are correlated and both were ranked second based on PC2. Interestingly, the Dendrogram grouped the Plantar GSR and Palmar GSR together, confirming the results in PC2. The multiple comparison of means reflected a significant difference between at rest and driving in the city and on the highway. In fact, there was a complete overlap between driving in the city and on the highway, as shown in Fig. 7, suggesting that the Plantar GSR is unable to capture changes in driving stress (or stress transition). Interestingly, the Plantar GSR was ranked third for detecting driving stress based on overall accuracy (66.28%), as shown in Table 2.

E. PALMAR GSR

Palmar GSR is the skin conductance measured from the hands. Typically, two electrodes are placed on the skin on the hand to capture changes in the skin's glands when producing ionic sweat [22]. Healy and Picard [11] reported that Palmar GSR is correlated with driving stress level. Palmar GSR was not considered in the original study [8], [11]. However, a recent study [5] showed that Palmar GSR was the biosignal most correlated with driving stress. As mentioned above, the unsupervised ML methods revealed that Plantar GSR and Palmar GSR are correlated, as shown in Fig. 4 and Fig. 5. The multiple comparison of means presented similar results as for the Plantar GSR, see Fig. 7, suggesting a significant difference between at rest and driving in the city and on the highway with an overlap between driving in the city and on the highway. This result shows that Palmar GSR is not sensitive for driving stress changes. The supervised ML methods ranked Palmar GSR second for detecting driving stress based on the overall accuracy (72.52%), as shown in Table 2.

F. RESP

RESP provides valuable information about breathing rate, which is considered an indicator of emotional states, including stress, arousal, and mental workload [23]. Typically, the sensor is integrated in a chest belt to measure the inhalation and exhalation by capturing the expansion and reduction of the belt size. The RESP signal has been examined previously to assess the affective state [23]–[25]. A previous

study [11] ranked the RESP signal third, while another study [5] ranked it second when compared with other biosignals when it comes to its correlation with driving stress. As mentioned above, PC1 showed that the RESP is correlated with the driving stress marker, as shown in Fig. 4. Moreover, the Dendrogram categorized the RESP near the driving stress marker; but placed it in a separate group (cf. Fig. 5). The multiple comparison of means presented similar results as for Plantar GSR, Palmar GSR, and EMG, as shown in Fig. 7. Note that ECG was able to detect the three driving zones while RESP detected driving in the city and in the highway as one zone. The overlap between driving in the city and on the highway suggests less sensitivity for detecting driving stress. The supervised ML methods ranked RESP fourth for detecting driving stress based on the overall accuracy (65.02%), as shown in Table 2.

G. HR

The variability in HR provides valuable information about the autonomic nervous system, which is considered a marker of driving stress [26]. The HR signal has been examined previously to assess the affective state [23]–[25]. Previous studies [5], [11] ranked the HR signal fourth when compared with other biosignals on the same dataset for detecting driving stress. It is important to note that the HR is extracted from the ECG signal by detecting the R peaks.

As mentioned above, the PCA analysis did not find a correlation between HR and the driving stress marker, as shown in Fig. 4. Moreover, the Dendrogram did not categorize HR near the driving stress marker. In fact, it was the farthest away (cf. Fig. 5). The multiple comparison of means revealed a complete overlap among all driving statuses, as shown in Fig. 7, suggesting that the HR is the least informative biosignal for a driving stress level assessment. The supervised ML methods ranked HR sixth (bottom of the rank) for detecting driving stress based on the overall accuracy (63.78%), as shown in Table 2. Interestingly, given that the HR is derived from ECGs, the ECG signals achieved a higher performance than the HR. This means that there is valuable information within the ECG waveforms that are correlated with driving stress. In other words, the RR intervals are not sufficient enough to capture driving stress. It is worth mentioning that the purpose of this analysis is to rank biosignals based on their abilities to detect driving stress.

H. WHICH BIOSIGNAL IS MOST CORRELATED WITH DRIVING STRESS?

To answer this question objectively, we compared previous studies using the same dataset. Table 3 compares the finding from our study with other studies that used the same dataset. Study 1 [11] and Study 2 [5] confirmed that in general, the GSR is the most informative biosignal for assessing driving stress levels. However, Study 2 reported Plantar GSR, while Study 1 reported Palmar GSR as optimal biosignals for detecting driving stress. Perhaps the reason behind getting

TABLE 3. Ranking biosignals based on their ability to assess driving stress using the Stress Recognition in Automobile Drivers (SRAD) database.

Rank	Study 1 [11] Supervised	Study 2 [5] Supervised	This work (PCA) Unsupervised	This work (Clustering) Unsupervised	This work (Classification) Supervised
1	Palmar GSR	Plantar GSR	ECG	ECG	ECG
2	EMG	RESP	RESP	RESP	Palmar GSR
3	RESP	Palmar GSR	Plantar GSR	EMG	Plantar GSR
4	HR	HR	Palmar GSR	Plantar GSR	RESP or EMG
5	-	EMG	HR	Palmar GSR	HR
6	-	-	EMG	HR	-

two different results is because Study 1 excluded Plantar GSR from the analysis.

Another contradiction between Study 1 and Study 2 pertains to the second most informative biosignal. Study 1 recommended HR as the second best (in combination with Palmar GSR), while Study 2 recommended RESP as second best. A reasoning for this contradiction is that Study 2 used wavelet-based features, while Study 1 used basic statistical features, such as mean and standard deviation. The most interesting finding from our analysis is that ECG was found to be the most informative biosignal among all biosignals after rigorous testing using unsupervised and supervised ML methods. Note that Study 1 and Study 2 did not include ECG in their analysis and that may be why this finding has not been previously reported.

We must question whether the overall accuracy achieved in our analysis is important or not. The ECG signals achieved an overall accuracy of 75.02% using only one feature extracted per second. This result outperformed that of a recent study [5], Study 2 (Table 3), which used 12 different features extracted from the Palmar GSR, Plantar GSR, and RESP, achieving 69% accuracy. It is worth mentioning that Study 2 used two-thirds of the dataset for training and one-third for testing, while in our study, a leave-one-out cross-validation was carried out. Study 2 used only five minutes from each segment, while our study considered the whole segment without removing any data points. Moreover, in the current paper, unsupervised and supervised techniques were used to rank the biosignals, while in Study 2, the findings were based on supervised learning. In addition, in this paper, both the raw data and extracted features were used to confirm the findings, while Study 2 relied on certain features only, which may have negatively impacted the overall driving stress detection accuracy.

We have attempted to objectively identify the optimal biosignal for assessing driving stress. The results consistently showed, using longitudinal, transitional and multicomparison analyses, that ECG is the optimal biosignal for a driving stress assessment. Note that, to our knowledge, this is the first study that has rigorously ranked multiple biosignals using different ML techniques. One must acknowledge several other pertinent factors that faced previous studies to reach to the same finding, such as the exclusion of some biosignal from the analysis, the rigorousness and clarity of the cross validation, the examination of different driving statuses such as (R1 vs. C1), (H1 vs. C1), (R1 vs. C2), (H2 vs. C2), and (R2 vs. C3),

examining the time-series data itself, and considering all data points in each driving zone.

V. PRACTICALITIES OF STUDY FINDINGS
The practicality of collecting ECGs and the long-term implications of doing so can only be speculated at this point. It is conceivable that this technology could help individuals to manage stress when driving, which in turn could reduce accident risk.

1) *ECG Sensor Placement*: Many attempts have been made in the past to monitor a driver's visual and cognitive distractions [27]. Yet, most of the techniques did not become a practical application (i.e., integrated in smart cars). Perhaps this is due to the practicality of placing the ECG sensor in the car and the questionable robustness of detecting stress during driving. We propose mounting the ECG electrodes on the steering wheel to monitor stress while driving. This approach seems to be more convenient and has a good chance of being accepted by The National Highway Traffic Safety Administration as it will not distract the driver.

2) *Improve Driving Experience*: When a smart car detects stress, with integrated ECG sensor, the car could recommend playing relaxing music or changing the vehicle temperature, as suggested in a previous study [28]. Adding an indication light feature on the dashboard could help with providing feedback to the drivers; perhaps increasing the sense of control they have over their driving state. For example, if the light on the dash board is green, indicating low stress, the driver may then in return feel calmer and more emotionally stable. In contrast, a red light would indicate that the driver is in a high-stress state and the smart car could recommend pulling over while playing calming music.

3) *Reducing Insurance Costs*: Maintaining a consistently low stress level while driving could help to build a solid credit history over the long-term, which consequently could lower insurance costs. Insurance companies usually offer discounts to policyholders who have not had any accidents or moving violations for a certain period of time. Similarly, increasing stress levels over a certain amount time may lead to further discounts.

VI. LIMITATIONS OF STUDY AND FUTURE WORK
One of the main limitations of the current study is the small sample size. The next step is to test the same methodology on a different database and see if the same results are achieved. However, the focus of the current study was not on improving the detection rate of driving stress. Rather, the focus was on ranking biosignals. Thus, different existing ML algorithms to rank the biosignals were used. One may question the difference between the best performance (75.02% with ECG) and worst performance (64.02% with HR) when it comes to detecting driving stress. At a first glance, this difference does not appear to be particularly significant. However, our aim was to identify which biosignal is more informative and

sensitive in detecting driving stress, for example, determining if ECG signal is more informative than HR, or vice versa. Even though the purpose of the present study was to rank the biosignals, the classification results outperformed other recent published works that focused on driving stress detection.

If the purpose of the current study was to improve the detection of driving stress using biosignals, specifically ECG, then more features would need to be extracted, or deep learning would need to be applied. Another approach that could be used to improve the detection of driving stress would be to combine multiple biosignals.

For formulating future studies, we recommend the collection of hemodynamics parameters (e.g., blood pressure levels and dilation of pupils) and to conduct psychological driving stress tests during the collection of biosignals. This will improve our understanding of driving stress levels and their correlation with biosignals changes.

VII. CONCLUSION

This paper has described the biosignal that best determines driving stress, by looking at biosignals such as ECG, EMG, Plantar GSR, Palmar GSR, HR, and RESP. Without prior training or knowledge, the INTENSE method (aggregating three ML methods: IPCA, CBC, and KMC) was able to detect a complex dynamic between different biosignals and the driving stress marker. There is agreement between all ML methods that a significant correlation exists between the driving stress marker and ECG. Moreover, when supervised ML methods were used, ECG scored the highest overall accuracy for detecting stress in different driving scenarios. One interesting result is that the HR did not correlate well with the driving stress marker. These findings are timely because more than ever, ECG-based wearable devices are being used on a large scale in different scientific fields. The results of this study can contribute to recommendations for the use of ECG signals as an informative means of measuring driving stress and anxiety. Using this newfound knowledge together with artificial intelligence, we can better use ECGs to not only monitor cardiac abnormalities, but also for assessing driving stress.

REFERENCES

[1] *Stress in America: The State of Our Nation, Stress in America Survey*, Amer. Psychol. Assoc., Washington, DC, USA, 2017.

[2] M. Kivimäki and A. Steptoe, "Effects of stress on the development and progression of cardiovascular disease," *Nature Rev. Cardiol.*, vol. 15, no. 4, pp. 215–229, Dec. 2017.

[3] J. M. Jakicic, K. K. Davis, R. J. Rogers, W. C. King, M. D. Marcus, D. Helsel, A. D. Rickman, A. S. Wahed, and S. H. Belle, "Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: The IDEA randomized clinical trial," *JAMA*, vol. 316, no. 11, p. 1161, Sep. 2016, doi: 10.1001/jama.2016.12858.

[4] M. Elgendi and C. Menon, "Assessing anxiety disorders using wearable devices: Challenges and future directions," *Brain Sci.*, vol. 9, no. 3, p. 50, Mar. 2019, doi: 10.3390/brainsci9030050.

[5] N. El Haouij, "Biosignals for driver's stress level assessment: Functional variable selection and fractal characterization," Ph.D. dissertation, Univ. Paris-Saclay, École nationale d'ingénieurs de Tunis (Tunisie), 2018.

[6] E. Smets, E. Rios Velazquez, G. Schiavone, I. Chakroun, E. D'Hondt, W. De Raedt, J. Cornelis, O. Janssens, S. Van Hoecke, S. Claes, I. Van Diest, and C. Van Hoof, "Large-scale wearable data reveal digital phenotypes for daily-life stress detection," *npj Digit. Med.*, vol. 1, no. 1, Dec. 2018, doi: 10.1038/s41746-018-0074-9.

[7] M. F. Rizwan, R. Farhad, F. Mashuk, F. Islam, and M. H. Imam, "Design of a biosignal based stress detection system using machine learning techniques," in *Proc. Int. Conf. Robotics,Electr. Signal Process. Techn. (ICREST)*, Dhaka, Bangladesh, Jan. 2019, pp. 364–368.

[8] J. A. Healey, "Wearable and automotive systems for affect recognition from physiology," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2000.

[9] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.

[10] P. Rowden, G. Matthews, B. Watson, and H. Biggs, "The relative impact of work-related stress, life stress and driving environment stress on driving outcomes," *Accident Anal. Prevention*, vol. 43, no. 4, pp. 1332–1340, Jul. 2011.

[11] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005, doi: 10.1109/TITS.2005.848368.

[12] M. Elgendi, "TERMA framework for biomedical signal analysis: An economic-inspired approach," *Biosensors*, vol. 6, no. 4, p. 55, Nov. 2016.

[13] M. Elgendi, "Characteristics of a highly cited article: A machine learning perspective," *IEEE Access*, vol. 7, pp. 87977–87986, Jul. 2019, doi: 10.1109/ACCESS.2019.2925965.

[14] M. Elgendi, "Optimal signal quality index for photoplethysmogram signals," *Bioengineering*, vol. 3, no. 4, p. 21, Sep. 2016, doi: 10.3390/bioengineering3040021.

[15] M. Elgendi, B. Eskofier, S. Dokos, and D. Abbott, "Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems," *PLoS ONE*, vol. 9, no. 1, Jan. 2014, Art. no. e84018, doi: 10.1371/journal.pone.0084018.

[16] R. Chaudhary, "Electrocardiogram comparison of stress recognition in automobile drivers on MATLAB," *Adv. Electron. Elect. Eng.*, vol. 3, no. 8, pp. 1007–1012, 2013.

[17] C. Karmakar, M. H. Imam, A. Khandoker, and M. Palaniswami, "Influence of psychological stress on QT interval," in *Computing in Cardiology*. Cambridge, MA, USA, 2014, pp. 1009–1012.

[18] K. A. Sidek and I. & Khalil, "Automobile driver recognition under different physiological conditions using the electrocardiogram," in *Proc. Comput. Cardiol.*, Cambridge, MA, USA, 2011, pp. 753–756.

[19] S. Jerritta, M. Murugappan, K. Wan, and S. Yaacob, "Emotion recognition from facial EMG signals using higher order statistics and principal component analysis," *J. Chin. Inst. Eng.*, vol. 37, no. 3, pp. 385–394, Jul. 2013.

[20] J.-W. Tan, A. O. Andrade, H. Li, S. Walter, D. Hrabal, S. Rukavina, K. Limbrecht-Ecklundt, H. Hoffman, and H. C. Traue, "Recognition of intensive valence and arousal affective states via facial electromyographic activity in young and senior adults," *PLOS ONE*, vol. 11, no. 1, Jan. 2016, Art. no. e0146691, doi: 10.1371/journal.pone.0146691.

[21] R. Luijcks, H. J. Hermens, L. Bodar, C. J. Vossen, J. V. Os, and R. Lousberg, "Experimentally induced stress validated by EMG activity," *PLoS ONE*, vol. 9, no. 4, Apr. 2014, Art. no. e95215, doi: 10.1371/journal.pone.0095215.

[22] W. Boucsein, *Electrodermal Activity*. Boston, MA, USA: Springer, 2012.

[23] A. H. Roscoe, "Assessing pilot workload. Why measure heart rate, HRV and respiration?" *Biol. Psychol.*, vol. 34, nos. 2–3, pp. 259–287, Nov. 1992.

[24] Y. Lin, H. Leng, G. Yang, and H. Cai, "An intelligent noninvasive sensor for driver pulse wave measurement," *IEEE Sensors J.*, vol. 7, no. 5, pp. 790–799, May 2007, doi: 10.1109/JSEN.2007.894923.

[25] H. J. Baek, H. B. Lee, J. S. Kim, J. M. Choi, K. K. Kim, and K. S. Park, "Nonintrusive biological signal monitoring in a car to evaluate a Driver's stress and health state," *Telemedicine e-Health*, vol. 15, no. 2, pp. 182–189, Mar. 2009, doi: 10.1089/tmj.2008.0090.

[26] J. F. Thayer, F. Åhs, M. Fredrikson, J. J. Sollers, and T. D. Wager, "A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health," *Neurosci. Biobehavioral Rev.*, vol. 36, no. 2, pp. 747–756, Feb. 2012, doi: 10.1016/j.neubiorev.2011.11.009.

[27] M. Kutila, M. Jokela, G. Markkula, and M. R. Rue, "Driver distraction detection with a camera vision system," in *Proc. IEEE Int. Conf. Image Process.*, San Antonio, TX, USA, Sep. 2007, p. VI-201, doi: 10.1109/ICIP.2007.4379556.

[28] J. Hernandez, D. McDuff, X. Benavides, J. Amores, P. Maes, and R. Picard, "AutoEmotive: bringing empathy to the driving experience to manage stress," in *Proc. Companion publication Designing Interact. Syst. DIS Companion*, 2014, pp. 53–56, doi: 10.1145/2598784.2602780.

MOHAMED ELGENDI (Senior Member, IEEE) is currently a Senior Postdoctoral Fellow at the UBC's Department of Obstetrics and Gynecology, an Adjunct Professor with the UBC's Department of Electrical and Computer Engineering, and a Senior Fellow at the Howard Brain Sciences Foundation. In addition to his 10+ years of experience in the field of data analysis, he received training on Big Data Analysis and Leadership in Education from MIT. His experience in the areas of digital health, data analysis, and visualization includes his work in global health with the PRE-EMPT Initiative (funded by the Bill and Melinda Gates Foundation), the Institute for Media Innovation at Nanyang Technological University (Singapore), and Alberta's Stollery Children's Hospital (Canada). He specializes in bridging the areas of engineering, computer science, psychology, and medicine for knowledge translation.

CARLO MENON (Senior Member, IEEE) received the Laurea degree in mechanical engineering and the Ph.D. degree in bio-robotics from the University of Padua, Padua, Italy, in 2001 and 2005, respectively.

He was a Research Fellow at the European Space Agency, The Netherlands, in 2005 and 2006, respectively. In 2007, he joined the School of Engineering Science, Simon Fraser University (SFU), Canada, as an Assistant Professor, and founded the Menrva Research Group, where he has been an Associate Professor, since 2012. He has published over 250 scientific works, including both journals and conference papers. He is a member of the American Institute of Aeronautics and Astronautics (AIAA), and the American Society of Mechanical Engineers. He received the International Astronautical Federation (IAF) Luigi G. Napolitano Award, Spain, in 2006, and the International Biomimetics Network for Industrial Sustainability (BIONIS) Award in Biomimetics, U.K., in 2007. In 2012, he received both the Career Investigator Award from the Michael Smith Foundation for Health Research (MSFHR) and the New Investigator Award from the Canadian Institutes of Health Research (CIHR). He is a member of the editorial board of the *Journal of Bionic Engineering* and an Associate Editor of *Bionics and Biomimetic* (Frontiers).

● ● ●