





Article

Towards a Contactless Stress Classification Using Thermal Imaging

Federica Gioia ^{1,2,*} , Alberto Greco ^{1,2} , Alejandro Luis Callara ^{1,2}  and Enzo Pasquale Scilingo ^{1,2} 

¹ Dipartimento di Ingegneria dell'Informazione, University of Pisa, 56122 Pisa, Italy; alberto.greco@unipi.it (A.G.); alejandro.callara@ing.unipi.it (A.L.C.); e.scilingo@ing.unipi.it (E.P.S.)

² Research Center "E. Piaggio", University of Pisa, 56122 Pisa, Italy

* Correspondence: federica.gioia@phd.unipi.it

Abstract: Thermal cameras capture the infrared radiation emitted from a body in a contactless manner and can provide an indirect estimation of the autonomic nervous system (ANS) dynamics through the regulation of the skin temperature. This study investigates the contribution given by thermal imaging for an effective automatic stress detection with the perspective of a contactless stress recognition system. To this aim, we recorded both ANS correlates (cardiac, electrodermal, and respiratory activity) and thermal images from 25 volunteers under acute stress induced by the Stroop test. We conducted a statistical analysis on the features extracted from each signal, and we implemented subject-independent classifications based on the support vector machine model with an embedded recursive feature elimination algorithm. Particularly, we trained three classifiers using different feature sets: the full set of features, only those derived from the peripheral autonomic correlates, and only those derived from the thermal images. Classification accuracy and feature selection results confirmed the relevant contribution provided by the thermal features in the acute stress detection task. Indeed, a combination of ANS correlates and thermal features achieved 97.37% of accuracy. Moreover, using only thermal features we could still successfully detect stress with an accuracy of 86.84% in a contact-free manner.



Citation: Gioia, F.; Greco, A.; Callara, A.L.; Scilingo, E.P. Towards a Contactless Stress Classification Using Thermal Imaging. *Sensors* **2022**, *22*, 976. <https://doi.org/10.3390/s22030976>

Academic Editor: Edward Sazonov

Received: 30 December 2021

Accepted: 25 January 2022

Published: 27 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: thermal imaging; stress detection; wearable systems; support vector machine; contactless

1. Introduction

Thermal imaging is currently taking hold in psychophysiology for its great advantage of measuring skin temperature noninvasively, ecologically, and contact free [1]. Many changes in skin temperature are regulated by the interplay between the parasympathetic (PNS) and sympathetic (SNS) branches of the autonomic nervous system, i.e., blood redistribution, perspiration, muscle activation, and metabolism. For instance, the activation of the SNS can alter the vasomotility of subcutaneous vessels (e.g., vasoconstriction) with consequent changes in the skin blood perfusion [2,3], or may control the sweat gland activation modifying skin perspiration [4]. Accordingly, monitoring skin temperature may provide information on SNS activity and its dysregulation induced by factors such as chronic or acute stress.

After a stressful stimulus, the sympathetic activation results in a series of rapid sympathetically driven adaptation changes of ANS activity, whose correlates are reflected by several peripheral signals. Among these, the breathing rate increases promoting oxygenation, the blood flow is redirected to large muscles through blood vessels dilation, and the heart rate and blood pressure increase, while digestive functions, urination, and feeling of hunger are inhibited [5]. Although stress is a physiological phenomenon that is helpful in many demanding situations, acute stress episodes could affect the human physical condition and health leading to several symptoms ranging from headache to heart attacks or arrhythmias [6–8].

Typically, stress is monitored by means of questionnaires, psychometric tests, and interviews. However, these methods provide subjective measures and can lead to a wrong assessment of the stress level when the subject is not completely honest or able to answer. Accordingly, many studies have attempted to infer the psychological state of a subject in a more reliable and unbiased manner by analyzing ANS peripheral correlates [9]. Among these, the electrodermal activity (EDA) could provide a reliable metric of stress, and it has been used as a ground-truth to compare the performance of other signals [10–13]. However, in view of a more comprehensive ANS assessment, a combination of autonomic measures, such as respiratory (RESP), cardiac, and EDA signals, is often adopted in the scientific literature, by using multiple wearable devices [14,15]. However, such an approach may affect the subject's natural behaviour both in laboratory settings and in real life scenarios mainly because of the presence of electrodes and cables. In this study, we contributed to the shift towards contactless stress monitoring by extending standard approaches to stress recognition with the aid of infrared thermography.

Thermal imaging unobtrusively monitors skin heat distribution mapping infrared (IR) radiation onto an image called a thermogram. So far, several physiological parameters have been extracted from the thermograms, such as breathing rate [16], cardiac pulse [17], and cutaneous blood perfusion [18]. In addition, information directly derived from the thermal time series has been analyzed in relation to mental states and emotional stimuli [19–23]. For example, the thermal variation of the nose tip has been used to discriminate stimuli with positive and negative valence [24], as well as in response to startling stimulus [25], but also to fear or stress [26–29]. Furthermore, thermal features from several facial regions have been used for statistical comparisons among different psychological states [1,30], and to classify emotional states [31,32]. In the stress-assessment context, stress elicited by the Stroop test was observed to induce an enhancement of blood flow in the forehead, supraorbital and frontal vessels, as indicated by different slopes of the blood volume signal obtained from the thermograms in baseline and stress condition [29].

We performed a systematic study to quantify the contribution of thermal imaging to a multisensor wearable system for stress recognition. In particular, we evaluated whether thermal features can improve the results obtained by a wearable system based on well-established physiological features extracted from EDA, HRV, and RESP signals in automatically and reliably recognizing stress. Moreover, we investigated the potential use of the contactless monitoring system in substitution of the wearable one. We followed a rigorous and robust methodology with the purpose of increasing the accuracy achieved by previous studies. At the same time, we included innovative facial regions (i.e., nasal septum, left and right forehead, and left and right chin) and thermal features (i.e., derivative of the mean temperature and its standard deviation). We compared the classification performance of a support vector machine with recursive feature elimination (SVM-RFE) classifier using different sets of physiological and thermal features. More specifically, we acquired physiological signals, i.e., electrocardiogram (ECG), EDA, and RESP along with thermal images from 25 healthy subjects performing a computerized/paced Stroop test. On this dataset, we conducted a preliminary statistical analysis comparing features extracted from the EDA, HRV, RESP, and thermal signals during resting and stressful conditions, with the aim of evaluating the effectiveness of our protocol in eliciting acute mental stress. Then, we evaluated whether thermal features improve stress/non-Stress classification accuracy when integrated with more standard physiological ones. Finally, we attempted to recognize stressful conditions by using thermal features alone. All of the details are reported in the next sections, organized as follows: in Section 2, we describe the protocol and acquisition devices, processing methods for each signal, feature extraction, statistical analysis, and classification algorithm. In Section 3 we show the experimental results and in Section 4 discuss our findings.

2. Materials and Methods

2.1. Study Population

A total of 25 healthy subjects (9 females, age = 27 ± 3.1 years) participated in this study. The experiments were approved by the “Bioethics Committee of the University of Pisa”

(n. 15/2019), and each volunteer signed the informed consent. Subjects were asked not to drink coffee the day of the experiment to limit potential confounds due to vasomotor effects (e.g., increased/decreased vasomotility) [33]. For the same reason, they were also asked to avoid the consumption of alcohol and tobacco starting on the day before the experiment.

2.2. Experimental Protocol

The experimental timeline is reported in Figure 1. It consisted of three 3 min long sessions preceded by a 10 min long acclimatization period.

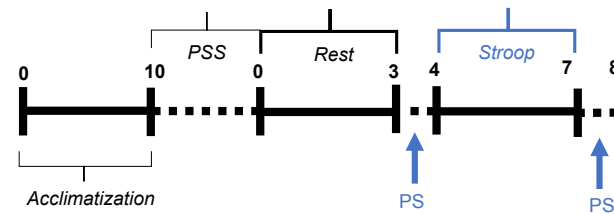


Figure 1. Protocol timeline. The dashed line indicates that the subject may take as much time as he/she needs to complete the task. PS indicates the moment when the subject is asked to report its perceived stress level.

The first session represented a baseline condition during which subjects were resting with their eyes open. The second session aimed to induce mental stress by means of the Stroop test [34]. This latter is a well-known stressor that requires the fast resolution of two incongruous stimuli causing cognitive interference. More specifically, incongruent color/semantic-meaning words are presented to a subject who has to answer based on the color of the word. Here, we developed a computerized paced version of the Stroop test, in which the name of a color was presented in the middle of the tablet screen every two seconds, and the subjects had to press the button corresponding to the tint of the word before a new stimulus appeared. We included a counter, keeping track of the consecutive success as a motivational stressor. During the task, a ticking sound was being played in the background to mark the passage of time. Additionally, for each wrong/missing answer the counter was reset while an acoustic buzzer alerted the subject. Finally, before and after the Stroop session, subjects were asked to report their level of perceived stress on a Likert scale from 0 (not at all) to 10 (very stressed). The self-assessed perceived stress level served as a reference for our analyses.

All the instructions were presented to the subjects on a tablet. Accordingly, interactions with the experimenter were limited, preventing speech artefacts in the data and standardizing the conditions for all subjects. Additionally, each subject filled out the Perceived Stress Scale (PSS) questionnaire. The PSS is a standard psychological test for measuring the perception of stress, as the degree to which situations in one's life are appraised as stressful [35].

The room temperature and humidity were monitored throughout the experiment by the EXTECH RH550 Humidity-Temperature Chart Recorder (Extech Instruments, Nashua, New Hampshire, U.S.) at a rate of 0.5 Hz. In particular, we allowed a maximal temperature variation within one experimental session of ± 1 °C and a humidity variation of $\pm 5\%$. Indeed, a stable environment was required to ensure reliable relative measures of facial skin temperature. Moreover, since these variables could trigger body thermoregulatory responses not related to emotional stress, temperature and humidity were always in a comfortable range of 22 °C to 28 °C and 40% to 56%, respectively [36]. Finally, the subject was placed in a seated position at a fixed distance of 1 m from the camera, at maximum distance from the window in order to avoid direct sunlight, and with no direct ventilation. Taking this into account, we could ensure reliable relative measures without adding the influence of ambient temperature and humidity on facial area temperature. In addition, subjects were asked to remove corrective eyewear during the experiment, as glass is opaque to infrared light.

2.3. Data Acquisition

2.3.1. Thermal Imaging

We acquired both thermal and RGB images of subjects' faces. The skin temperature distribution was unobtrusively acquired through a FLIR T640 thermal camera (Teledyne FLIR LLC, Wilsonville, Oregon, U.S.), at a rate of 5 Hz. The camera had a resolution of 640×480 pixels, 24.6 mm lens, NETD < 0.04 mK @ $+30$ °C and spectral range of 7.8–14 μ m (long-wave infrared—LWIR). Additionally, standard RGB images were acquired at a rate of 30Hz using a Logitech HD Webcam C270 (Logitech International S.A., Lausanne, Switzerland), 1280×720 pixels, vertically aligned and fixed to the thermal camera.

2.3.2. Physiological Signals

We recorded EDA, ECG, and RESP using a BIOPAC MP 150 system (Biopac Systems, Inc., Goleta, CA, USA), at a sampling rate of 500 Hz. EDA was recorded placing two electrodes on the distal phalanx of the index and ring fingers of the non-dominant hand. ECG was recorded with 3 electrodes positioned below the right and left clavicle, and on the lower left chest. Finally, RESP was monitored with a piezoelectric chest belt.

2.4. Data Processing

2.4.1. Thermal Processing

We used thermal imaging to monitor the variation in facial temperature distribution during the experiment and RGB frames to identify anatomical landmarks of interest for the analysis of IR images. Accordingly, RGB and IR acquisitions were synchronized after downsampling the RGB video to 5 Hz (i.e., IR imaging sampling rate). Then, RGB and IR frames were spatially registered with an affine-2D transformation matrix estimated from a set of manually selected fiducial points in the first frame of the RGB and IR datasets, such as periauricular points, the corner of the eyes and the mouth. After RGB-IR spatiotemporal registration, we defined 14 regions of interest (ROIs): i.e., nose tip, forehead (central, right and left), cheek (right and left), nasal septum, chin (central, right, left), periorbital (right and left), and maxillary (right and left). A summary of these ROIs along with their corresponding location is represented in Figure 2.

For each subject, ROIs were automatically found with a two-step procedure. First, the ROI centres were identified through the intersection of lines connecting specific facial landmarks detected in the RGB frames with the Yuval–Nirkin algorithm [37]. Then, each ROI was designed to have a proportional size to the total number of pixels of the face. Of note, the number of face pixels was automatically segmented by an intensity-landmark-based algorithm purposely developed for this study. Specifically, we first separated the face from the neck and the torso using the previously estimated face contour landmarks. Then, we excluded hair and background by segmenting IR maps with a threshold-based segmentation on the pixel temperature in range (30 °C $< T_p < 38$ °C). Finally, we tracked the centres of the ROIs throughout the frames with a Matlab PointTracker object based on Kanade–Lucas–Tomasi feature-tracking algorithm [38], to properly control for potential undesired subject's movement. For each frame, we evaluated the median temperature value of each ROI, obtaining 14 thermal signals per subject. High-frequency noise was filtered out by a moving median filter and the outliers (temperature values exceeding 3 standard deviations) were replaced by their nearest value. Then, from each thermal signal, we computed 4 features: the mean (**Mean**), the standard deviation (**Std**), and the mean and the standard deviation of the signal's derivative (**DMean**, **Dstd**). Of note, the thermal camera automatic non-uniformity correction (NUC) was performed before the beginning of the experiment for reliability purposes. Finally, NUC was disabled during the acquisition as we observed that it corrupted the signals by introducing abrupt thermal changes [30].

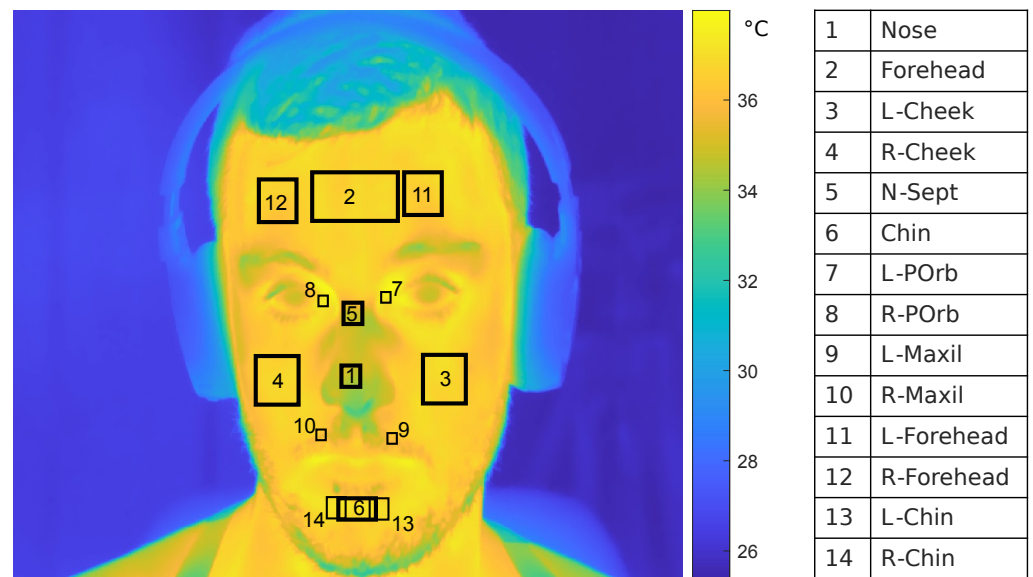


Figure 2. Facial ROIs' names and corresponding locations.

2.4.2. EDA Processing

The EDA signal reflects changes in the skin conductance induced by sweat glands' activity, which is controlled by the sympathetic branch of the ANS [4]. Therefore, EDA signal processing is a powerful tool to quantify the SNS dynamics. More specifically, the EDA signal can be decomposed into a tonic and a phasic component, which are characterized by different time scales. The tonic is a slow-varying component whose spectrum is below 0.05 Hz. Conversely, the phasic reflects short-term and event-related responses. Here, we used the cvxEDA model to extract tonic and phasic components from the raw EDA signal. Such a model exploits a rigorous framework integrating Bayesian statistics, convex optimization, and sparsity, to decompose the EDA signals without the need of pre- and post-processing steps [39]. Indeed, one of the outputs of the cvxEDA algorithm is the white Gaussian noise term, which incorporates model prediction errors as well as measurement errors and artefacts. Afterwards, for each experimental session we extracted several features related to the SNS activity starting from the tonic and phasic components: i.e., the mean tonic value (**TonicMean**) and the standard deviation of the tonic (**TonicStd**) averaged within non-overlapped 20 s long time windows; the maximum peak (**PksMax**), the mean of the phasic component (**PhasicMean**), the standard deviation of the phasic component (**PhasicStd**), the number of peaks (**NPks**), the sum of peaks (**PksSum**) averaged within non-overlapped 5 s long time windows, and the **EDAsymp** (power of the SC spectrum in the range of 0.04 Hz–0.25 Hz) [40], within the entire *Rest* and *Stroop* sessions.

2.4.3. HRV Processing

ECG recordings were analyzed with Kubios HRV [41]. The interbeat (RR) series were extracted from the ECG using the well-known Pan–Tompkins algorithm [42]. Algorithm-related artefacts were removed by applying the cubic spline interpolation method. Then, the obtained RR time series were resampled at 4 Hz to derive the heart rate variability (HRV) signals [43]. Starting from the HRV time series, several features were derived in the time, frequency, and nonlinear domain for both the *Rest* and *Stroop* session. Specifically, for each condition, we estimated in 3 min long segments the mean value of HRV signal (**mean HRV**), the standard deviation of HRV (**std HRV**), the square root of the mean squared differences of successive normal-to-normal (NN) intervals (i.e., **RMSSD**), the percentage number of pairs of adjacent NN intervals differing by more than 50 ms (**pNN50**), the power expressed as a percentage of total power in the low-frequency (**LF**, 0.04–0.15 Hz) and high frequency (**HF**, 0.15–0.40 Hz) ranges, the ratio of LF to HF power (**LF/HF ratio**), the minor

(SD1) and major (SD2) axis of the ellipse that best fits the Poincaré plot of RR intervals, and the sample entropy (**SamplEn**).

2.4.4. RESP Processing

The RESP signal was used as a further index of sympathovagal activity. Specifically, we estimated the mean frequency of respiration (**RESP freq**) for each condition and used it as a potential discriminating feature among conditions.

2.5. Exploratory Statistical Analysis

We performed an exploratory statistical analysis to assess at the group level how *Rest* and *Stroop* sessions differed in terms of ANS correlates. Five subjects were excluded from the analysis on EDA features and one from the thermal analysis due to the presence of artefacts related to excessive and frequent movement. Accordingly, these 6 subjects were excluded from the subsequent classification analyses. A non-parametric Wilcoxon sign-rank test was performed for each EDA, HRV, RESP, and thermal feature between *Rest* and *Stroop* sessions ($\alpha = 0.05$). We controlled false discovery rate through the Benjamini–Hochberg (FDR-BH) correction for multiple hypothesis testing [44]. To evaluate the psychometric changes induced by the *Stroop* session, the same test was performed on the self-assessed stress scores recorded before and after the task.

2.6. Classification—Stress and Rest Recognition

We explored the contribution of thermal features in the automatic recognition of *Rest* and *Stroop* sessions. In particular, we compared three SVM models for stress classification. More specifically, we performed three supervised binary classifications using three different sets of features: a full feature-set (*Full set*) composed of features from HRV, EDA, RESP, and thermal imaging, a set without the thermal features (*No-Thermo set*) and a set exclusively made of thermal features (*Thermo set*).

Given the large number of features in our dataset compared to the number of observations, it is crucial to reduce the dimensionality of the classification problem by applying a feature selection strategy (FS). To this aim, we trained an SVM-RFE [45] model that implements an embedded FS based on recursive feature elimination (RFE) strategy. More specifically, SVM-RFE is an application of RFE using the weight magnitude as a ranking criterion and iteratively removing the feature that has the least influence on the SVM weight-vector magnitude. It is worthwhile to note that highly correlated features can bring wrong estimations of features importance in feature selection algorithms. In fact, due to the so-called correlation bias, correlated features can receive underestimated ranking criteria, with a consequently higher risk to be removed [46]. Accordingly, in our study, we adopted the SVM-RFE algorithm that implements a correlation bias correction (CBR) (see [45]). In this way, the ranking criteria is corrected for collinearity by the CBR. Embedded methods as the one implemented here allow not only maximizing the classification accuracy but also facilitating data understanding thus reducing the risk of overfitting. They are defined as embedded because the search of the optimal subset is built into the classifier construction [47]. However, if the FS is embedded into the classifier construction, then the initial number of features in the input dataset should not be comparable with or even greater than the number of observations, otherwise, this would affect the correct estimation of the SVM model weights on which the RFE is based.

For the aforementioned reason, here, we implemented a two-stage feature selection algorithm [48]: first, we reduced the initial number of features by performing a preliminary filter-type feature selection based on the correlation between each feature and the dependent variable (the class label) [49]; secondly, we applied the SVM-RFE model on the pre-filtered dataset. More specifically, in the first stage, for each feature we computed the correlation with the corresponding class (labels: 1 = *Stroop*, 2 = *Rest*) and selected the features for which correlation was significant (p value ≤ 0.05).

The classifiers were implemented in Matlab using the LIBSVM library. We selected a nonlinear ν SVM-RFE with a radial basis function (RBF) kernel, and default parameters values. As for the out-of-sample testing, we used a leave one subject out (LOSO) cross-validation to avoid bias in the accuracy level estimation. Specifically, at each iteration, all data from all subjects were used as a training set except one subject, which was used as a test set. The choice of the subject to be excluded from the training set at each iteration was not random. Indeed, all subjects were used as testing set at some point in the iterations.

3. Results

An example of processed thermal signals of a sample subject for each ROI is reported in Figure 3. In Figure 4, we also report an example of the resulting signals of the cvxEDA decomposition algorithm for a sample subject.

PSS tests showed that more than 75% of the subjects were inclined to moderately feel stress. More specifically, 4 subjects out of 25 exhibited low stress, 19 moderate stress, and 2 high stress.

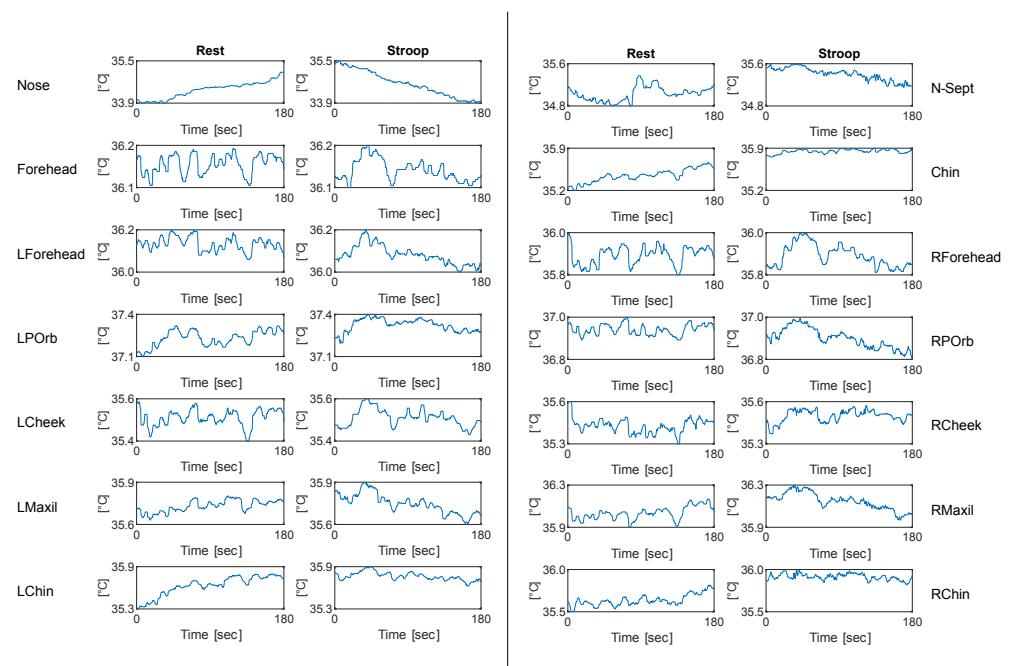


Figure 3. Thermal signals of one sample subject for each ROI in both rest and stress conditions. The temperature range of the plots was set equal for both conditions in each ROI, to allow visual comparison.

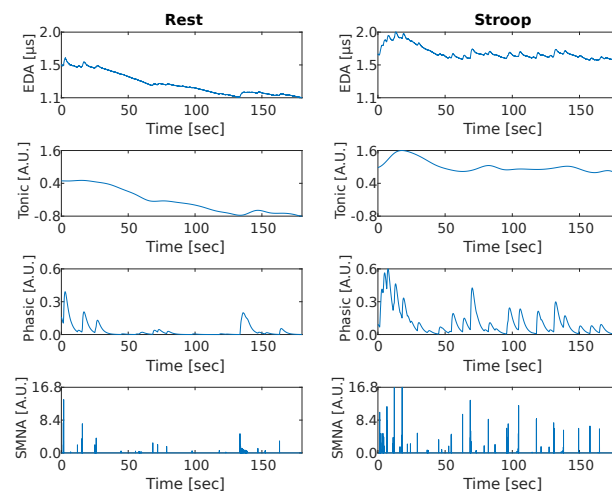


Figure 4. Application of the cvxEDA algorithm decomposition to the EDA signal of one sample subject in both rest and stress conditions.

3.1. Exploratory Statistical Analyses

The Wilcoxon test on the reported level of perceived stress indicated that subjects were more stressed after the task compared to before (median \pm mad = 2 ± 1.12 vs. median \pm mad = 6 ± 1.87 , $p < 0.001$). Concerning the physiological data analysis, most of the features extracted from autonomic signals confirmed the expected significant difference between *Stroop* and *Rest* sessions. Specifically, except for **EDASymp**, all EDA features significantly differed between the two states, as shown in Figure 5. The significant increase of the EDA features indicates a higher activation of the SNS during the task than during rest at both tonic and phasic level.

Moreover, all the HRV related features showed a significant variation during *Stroop* compared to *Rest* sessions, with the only exception of **stdHRV** and **SD2** (Figure 6). Specifically, **MeanHRV** and the **LF/HF ratio** increased during the *Stroop* compared to the *Rest* session. On the other hand, the remaining features decreased during the *Stroop* session. Among these, a variation of interest is the reduction in HF power during the *Stroop* session, considering its relation to PNS activity. Conversely, the respiratory frequency increased significantly during the *Stroop* compared to the *Rest* session, as shown in Figure 7.

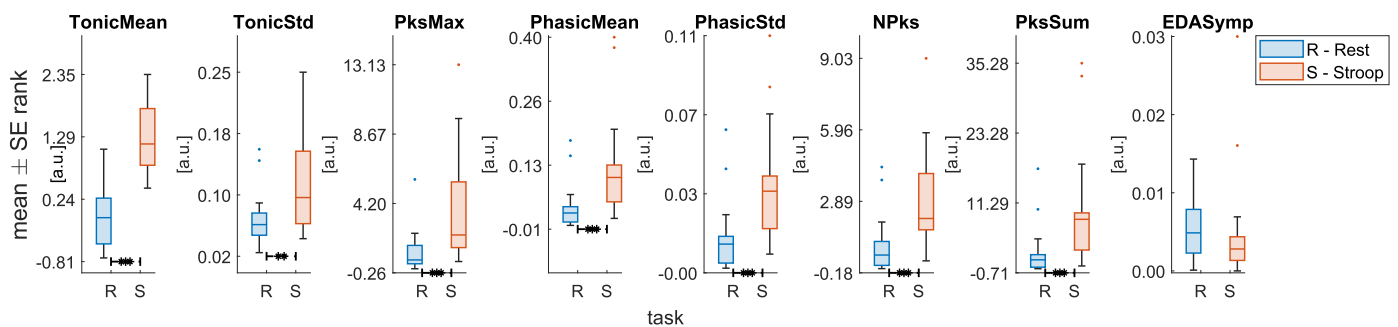


Figure 5. Statistical comparison for each of the EDA features extracted. Significant differences between *Rest* and *Stroop* sessions after the Wilcoxon signed rank test with FDR-BH correction are highlighted with an asterisk (* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$).

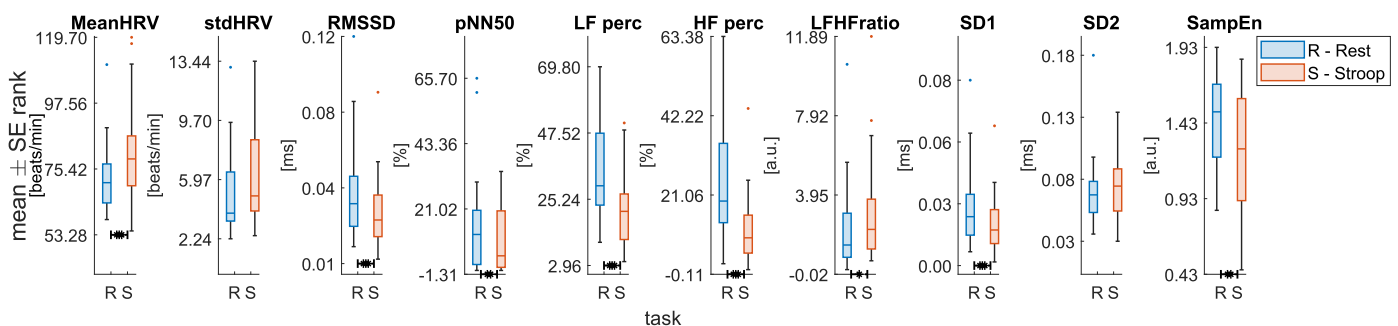


Figure 6. Statistical comparison for each of the HRV features extracted. Significant differences between *Rest* and *Stroop* sessions after the Wilcoxon signed rank test with FDR-BH correction are highlighted with an asterisk (* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$).

The results of the statistical analysis performed on the thermal features are shown in Figure 8. Among the considered ROIs, L-Forehead, Chin, L-Chin, and R-Chin did not show any significant alteration between sessions. Conversely, the remaining ROIs, except for N-Sept, showed a significant increase in the **Std** during the *Stroop* session. On the other hand **Mean** and **DMean** of Nose, N-Septum, and R-POrb decreased significantly during the *Stroop* session. Similarly, **DMean** of L-POrb showed a significant decrease during the *Stroop* session. On the contrary, **Mean** of L-Cheek increased significantly during the *Stroop* compared to the *Rest* session. Finally, **DStd** did not show any significant variation between the tasks.

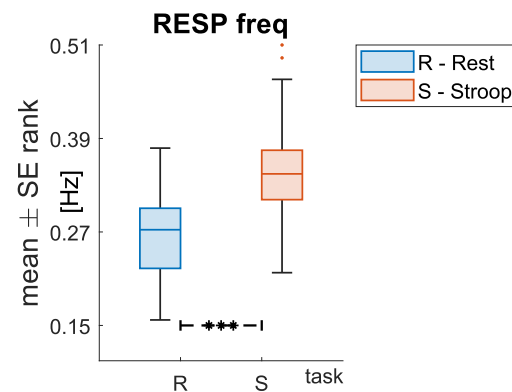


Figure 7. Statistical comparison for the respiratory feature extracted. Significant differences between *Rest* and *Stroop* sessions after the Wilcoxon signed rank test with FDR-BH correction are highlighted with an asterisk (* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$).

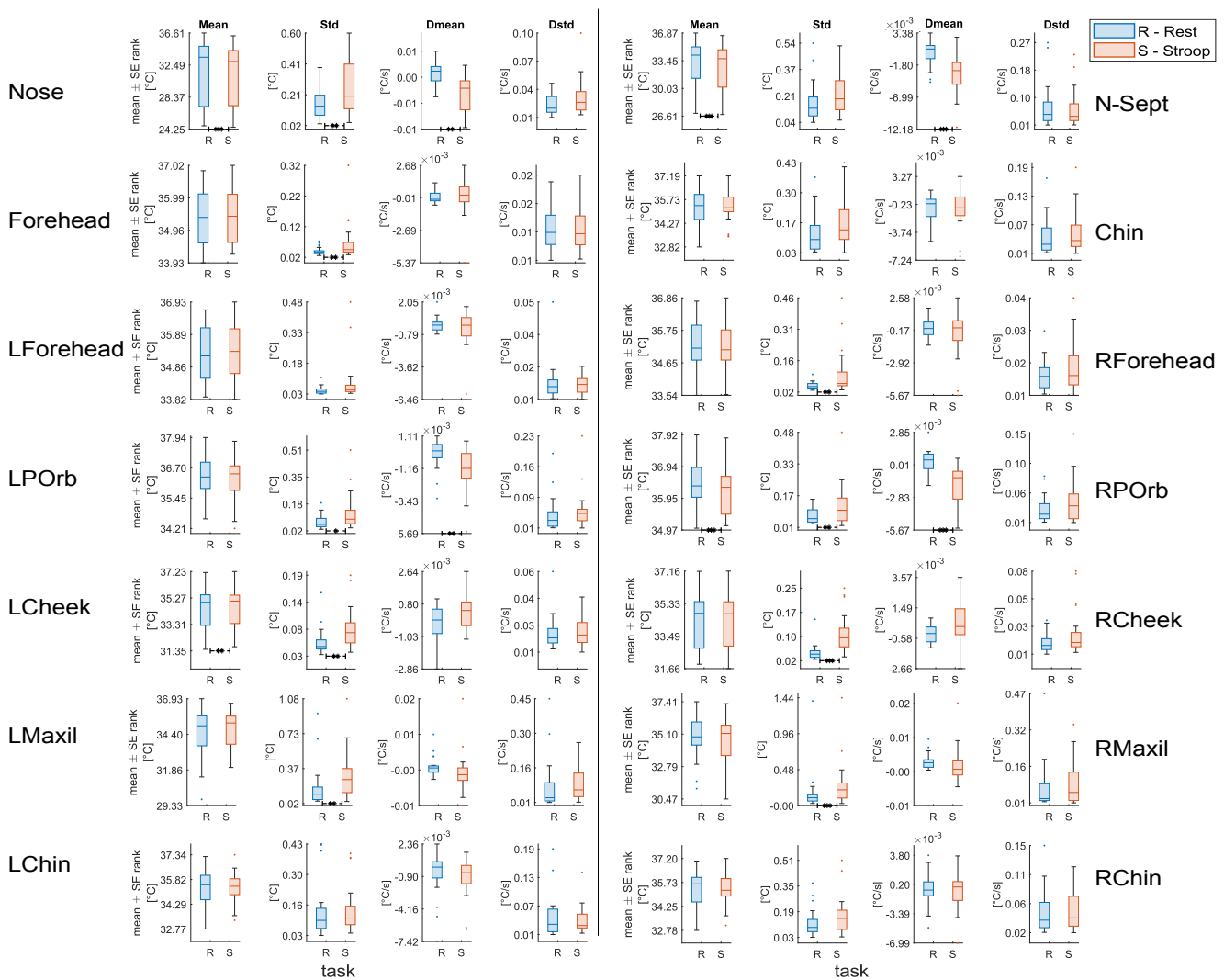


Figure 8. Statistical comparison for each of the thermal features extracted. Significant differences between *Rest* and *Stroop* sessions after the Wilcoxon signed rank test with FDR-BH correction are highlighted with an asterisk (* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$).

3.2. Classification

After the preliminary filter-type feature selection, *Full Set* dimensionality was reduced to 20, *Thermo Set* to 9, and *No-Thermo Set* to 11, as reported in Table 1. On these three different

sets, we compared the outcomes of the SVM-RFE classifier. Specifically, for each set, we evaluated the classification accuracy by means of LOSO cross-validation. The average accuracy at each cycle is reported in Figure 9. In particular, we observed a maximum accuracy of 97.37% for the *Full set* and 94.74% for the *No-Thermo set*, whereas for the *Thermo set* we observed a maximum accuracy of 86.84%. In particular, the peak of the accuracy trend for the *Full Set* was obtained with two thermal features and one EDA feature. The accuracy peak in the *Thermo Set* was obtained when the first five most relevant features, according to the RFE criterion, were considered. In the *NoThermo Set*, the maximum the accuracy was reached when four features were considered (three EDA features and the respiratory frequency).

In Table 1, for each set, we report the results of the filter-type feature selection procedure, along with the feature ranking according to RFE. We observed that within the *Full set*, nine out of twenty features were thermal ones and two of them (i.e., **RPOrb DMean**, **N-Sept DMean**) ranked among the first three features. In the *No-Thermo set*, all EDA features except **EDAsymp** were selected, together with **RESP freq**, **mean HRV**, and **std HRV**. Notably, the **TonicMean** was ranked as first in both *Full set* and *No-Thermo set*. In the *Thermo set*, **RCheek Std** was the first ranked feature, with an accuracy higher than 80%. The same accuracy was reached in the other two feature sets when considering only **TonicMean** from the EDA signal. Finally, in the *Thermo set*, several ROIs were selected (Table 1). In particular, only features based on signal derivative mean (i.e., **DMean**) and signal standard deviation (i.e., **Std**) were included.

Table 1. Feature ranking according to the RFE criterion.

Rank	Full Set	Rank	Thermo Set	Rank	No-Thermo Set
1	TonicMean	1	RCheek Std	1	TonicMean
2	RPOrb Dmean	2	RPOrb Dmean	2	PhasicMean
3	N-Sept Dmean	3	Nose DMean	3	RESP freq
4	PhasicMean	4	LCheek Std	4	PksSum
5	PksSum	5	N-Sept DMean	5	TonicStd
6	PhasicStd	6	RPOrb Std	6	NPks
7	TonicStd	7	LPOrb DMean	7	PhasicStd
8	Npks	8	RForehead Std	8	SampEn
9	RESP freq	9	Forehead Std	9	PksMax
10	PksMax			10	std HRV
11	RCheek Std			11	mean HRV
12	Nose DMean				
13	SampEn				
14	std HRV				
15	RPOrb Std				
16	LPOrb DMean				
17	RForehead Std				
18	LCheek Std				
19	mean HRV				
20	Forehead Std				

For each feature set, we report the confusion matrices related to the subset that achieved the maximum accuracy (Table 2). For the *Thermo set*, we observed a percentage of

false negatives (error type II) of 15.79%, lower values were found in the *Full set* and in the *No-Thermo set*.

Table 2. Confusion matrices. N-S: non-stress; S: stress. The percentage of true positives and true negatives of each class is represented in the diagonal of each table.

		Predicted Classes					
		S		N-S		S	
Actual Classes	S	94.74%	5.26%	84.21%	15.79%	100%	0%
	N-S	0%	100%	10.53%	89.47%	10.53%	89.47%
		<i>Full Set</i>		<i>Thermo Set</i>		<i>No-Thermo Set</i>	

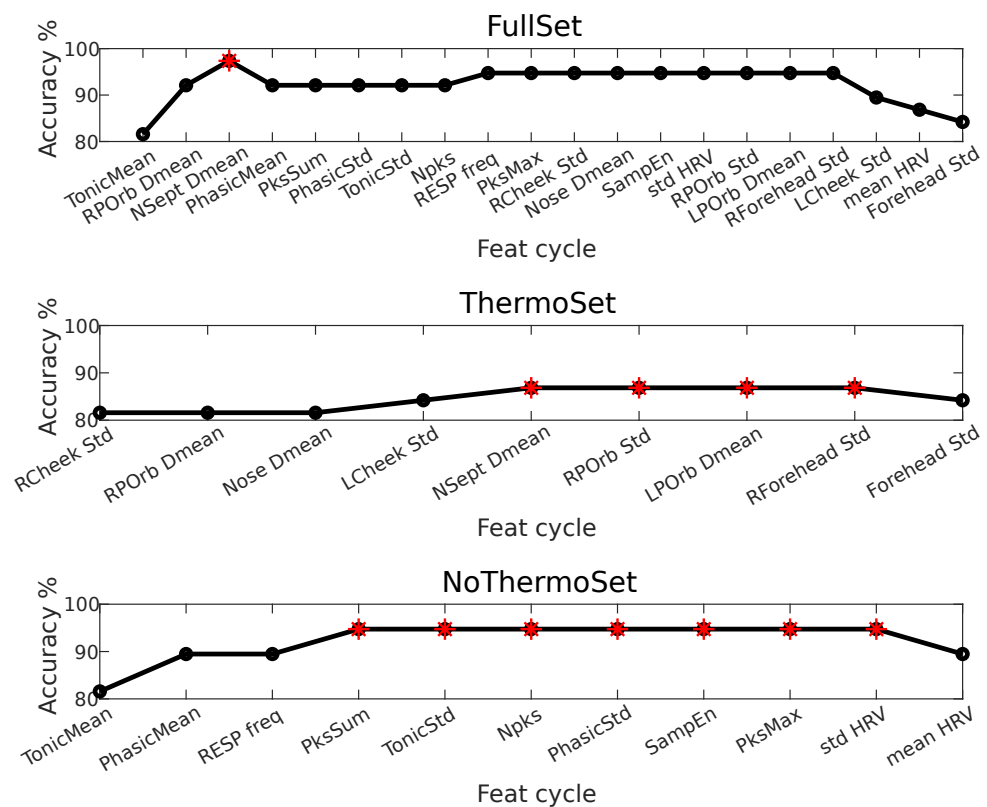


Figure 9. Accuracy trend of stress/non-stress recognition problem as a function of the selected features. The features are ranked according to the SVM-RFE criterion. Maximum accuracy is highlighted by red asterisks.

4. Discussion

In this study, we explored the possibility of using the thermal camera for automatic acute stress detection either in combination with other typical physiological signals or alone. In this latter condition, we made a relevant step towards a reliable contactless subject-independent stress recognition system exclusively based on thermal features. This might contribute to overcoming issues given by multiple wearable devices and support the shift towards contactless stress monitoring.

We designed a computerized and paced version of the Stroop test to mimic a more realistic stressful situation [50,51]. This allowed to reduce the interaction of the subjects with the experimenter, to standardize the external stimuli for all the participants, and to prevent speech artefacts in the signals. On the other hand, the use of well-established measures for stress recognition allowed us to confirm the effectiveness of the stimulation protocol. Here, we observed a significant increase in self-assessed perceived stress levels

after the *Stroop* session, as expected. Furthermore, the group-level preliminary statistical analyses on physiological signals supported this observation, as significant variations of most of the physiological parameters taken into consideration were observed. Among the HRV related features, we observed a decrease in **RMSSD**, **pNN50**, and **HF** which are likely to reflect a decrease in the PNS activity [52,53]. Moreover, we observed an increase of the **LF/HF ratio** which was previously associated with psychological stress due to changes in the sympathovagal balance [54]. To study the effect on the SNS activity, we based on the analysis of the EDA. Indeed, many of these features were found to be particularly suitable for stress monitoring [10,55]. In our study, we observed an increase in the mean value of the tonic component (**TonicMean**) during the *Stroop* session, which is likely to reflect the overall increase of the subject's arousal level. Furthermore, the increase of phasic component features indicated a more frequent and stimulus-related SNS activation, which may be related to the repetitive task-induced request for attention. Finally, we observed a higher respiratory frequency during the *Stroop* session, possibly reflecting sustained attention and mental stress [56].

The preliminary statistical analysis performed on thermal features allowed us to distinguish among rest and stress conditions as well, with the main advantage of providing the measures of interest in a contactless fashion. Stressful conditions were found to trigger a sympathetically driven vasoconstriction of the blood vessels in the skin [19,20,27,57], which is reflected by a drop in the temperature of the tip of the nose [22,24,58]. In our study, we observed such temperature drop as well. Additionally, our results extend such significant variations also to other facial regions, i.e., forehead, cheeks, chin, periorbital, and maxillary areas. In particular, we found an increase in the mean temperature of the forehead during *Stroop*, in agreement with [29], and which may be associated with the activation of the corrugator muscle in response to stress [19]. Moreover, we observed a significant temperature increase in the left cheek, which could be due to the well-known blushing phenomenon due to embarrassment [59] and possibly associated with social stress. It is worthwhile noting that some of these ROIs did not show a coherent thermal pattern throughout the literature [60]. In this view, although further investigation will be needed, we can confirm that several of these ROIs were informative for stress/non-stress recognition by means of the SVM-RFE algorithm used in this work.

We evaluated the contribution of thermal features in a stress recognition task by comparing the performance of a classifier trained with and without thermal features (i.e., *Full set* vs. *No-Thermo set*). Furthermore, we evaluated the ability of the thermal features alone to recognize stress (i.e., *Thermo set*).

We built a classification of stress at the single-subject level [61], avoiding possible bias introduced by other validation strategies such as the K-fold ones (when data from the same subject is presented both in the training and validation sets). Our results showed a significant contribution given by the features derived by the thermal signal which increased the accuracy from 94.74% to 97.37% (*Full Set*). Moreover, the peak of accuracy was obtained with only three features, two of which extracted by the thermal signal (i.e., **DMean** of RPOrb and N-Sept).

Notably, the classification exclusively based on thermal features achieved a maximum accuracy of 86.84%. Although this value was lower than the one obtained with wearable devices, our result suggests that stress recognition by means of contactless thermal imaging can still achieve good results. In this view, it is worthwhile noting that stress recognition based only on thermal imaging would have the advantages of being more comfortable with respect to intrusive contact-based systems.

To minimize the risk of overfitting due to the limited number of observations, we decreased the complexity of the model, maximizing the classification accuracy. In particular, we implemented a two-stage feature selection (FS) strategy to reduce the dimensionality of the classification problem. First, we reduced the number of features by performing a preliminary filter-type feature selection based on the correlation between each feature and the dependent variable (i.e., the class label) [49]. Moreover, we applied the SVM-

RFE+CBR algorithm on the reduced dataset [45]. This is an embedded FS that not only reduces the risk of overfitting, but also implements a correlation bias reduction strategy. Moreover, it has been proved that embedded FS methods outperform the filter and wrapper approach, scoring features' importance based on the output of the predictive model [62–64]. Finally, we evaluated generalization capabilities through leave-one-subject-out (LOSO) cross-validation strategy which made our results unbiased and subject independent.

The feature selection step allowed to highlight the most important features for optimal classification. Many thermal features were selected for the classification when using the *Full Set*. Notably, 9 features out of 20 were thermal ones. It is worthwhile noting that among the thermal features only the standard deviation **Std** and the mean of the derivative **DMean** were selected (both of them being measures of thermal variation). Specifically, **Std** is a measure of the amount of variation in temperature, while **DMean** is an index of the velocity of temperature variations. This result, in line with other studies [65], suggests that absolute temperature is of less interest compared to its relative changes. Other previous studies have tested the performance of combined or singleton wearable and contactless systems [66–68], with lower accuracy [67] or showing some limitations in the methodological approach [66,68]. Whereas the authors in [68] based the classification on respiratory features extracted from the thermograms, regardless of strictly-thermal signatures, another study [66] employed a decision tree classifier with LOSO validation to investigate whether feature fusion of the physiological and thermal modalities could further improve the acute stress detection rate. Despite their good accuracy, our processing and classification pipeline was able to outperform the decision tree classifier.

A potential limitation of psychophysiological studies involving thermal imaging is related to the confounding factors due to the surrounding environment. Infrared technology is highly sensitive to reflected radiation from nearby objects. In this work, acquisitions were performed in a controlled environment (e.g., no direct ventilation on the subjects and reasonable distance from the window). Furthermore, we prevented the subject's thermoregulation processes by controlling for room temperature and humidity to be steady and at a comfortable level [69]. Indeed, these processes could hide the small thermal variation due to the psychological factors under consideration. In addition, we have ensured an acclimatization period prior to the experiment of 10 min, as suggested in [70], to allow for skin temperature patterns stabilization [36,71]. Accordingly, we can assume that the observed physiological changes are due to the SNS/PNS balance in response to emotional stress rather than to other phenomena. Finally, another limitation is given by the movement constraints that are needed to ensure low ROI tracking errors and therefore low noise in the thermal signals. Some studies use a chin rest to avoid sudden head rotation or direction change [72]. Here, we asked the subject to remain still during the acquisitions. Afterwards, we visually inspected the signals and removed subjects with irreducible artefacts. However, further studies will try to investigate a method able to mitigate these limitations.

It is worthwhile to mention that different kinds of stimuli may induce stress (e.g., emotional, physical). In this work, we used only one type of mental stressor, obtained by means of the well-known Stroop test. However, future studies will consider more complex protocols (e.g., by including different stressors) to extend the results to different kinds of stress. Other insights could include thermal feature extraction also in the frequency or non-linear domains, possibly improving the performances of the classifier. Finally, in the perspective of extending these results to clinical uses, the developed system could be tested also with pathological subjects.

Author Contributions: Conceptualization, F.G., E.P.S., and A.G.; methodology, A.G. and F.G.; software, F.G.; validation, A.G. and F.G.; formal analysis, F.G.; investigation, F.G. and A.G.; resources, F.G. and A.G.; data curation, F.G.; writing—original draft preparation, F.G. and A.L.C.; writing—review and editing, all authors; visualization, A.L.C. and F.G.; supervision, A.L.C., E.P.S. and A.G.; project administration, E.P.S. and A.G.; funding acquisition, E.P.S. and A.G. All authors have read and agreed to the published version of the manuscript.

Funding: The research leading to these results has received partial funding from the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence). This research has received partial funding from European Union Horizon 2020 Programme under grant agreement n 824153 of the project “POTION-Promoting Social Interaction through Emotional Body Odours”.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Bioethics Committee of the University of Pisa Review No. 14/2019, 3 May 2019.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data are not publicly available due to privacy restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ioannou, S.; Gallese, V.; Merla, A. Thermal infrared imaging in psychophysiology: Potentialities and limits. *Psychophysiology* **2014**, *51*, 951–963. [[CrossRef](#)] [[PubMed](#)]
- Hall, J.E.; Hall, M.E. *Guyton and Hall Textbook of Medical Physiology E-Book*; Elsevier Health Sciences: New York, NY, USA, 2020.
- Wallin, B.G. Sympathetic nerve activity underlying electrodermal and cardiovascular reactions in man. *Psychophysiology* **1981**, *18*, 470–476. [[CrossRef](#)] [[PubMed](#)]
- Critchley, H.D. Electrodermal responses: What happens in the brain. *Neuroscientist* **2002**, *8*, 132–142. [[CrossRef](#)] [[PubMed](#)]
- McCorry, L.K. Physiology of the autonomic nervous system. *Am. J. Pharm. Educ.* **2007**, *71*, 78. [[CrossRef](#)]
- Greene, S.; Thapliyal, H.; Caban-Holt, A. A Survey of Affective Computing for Stress Detection: Evaluating technologies in stress detection for better health. *IEEE Consum. Electron. Mag.* **2016**, *5*, 44–56. [[CrossRef](#)]
- Stress revisited: a critical evaluation of the stress concept. *Neurosci. Biobehav. Rev.* **2011**, *35*, 1291–1301 [[CrossRef](#)]
- Krantz, D.S.; Whittaker, K.S.; Sheps, D.S. Psychosocial risk factors for coronary heart disease: Pathophysiologic mechanisms. In *Heart and Mind: The Practice of Cardiac Psychology*; American Psychological Association: Washington, DC, USA, 2011.
- Mohino-Herranz, I.; Gil-Pita, R.; Ferreira, J.; Rosa-Zurera, M.; Seoane, F. Assessment of mental, emotional and physical stress through analysis of physiological signals using smartphones. *Sensors* **2015**, *15*, 25607–25627. [[CrossRef](#)]
- Greco, A.; Valenza, G.; Lázaro, J.; Garzón-Rey, J.M.; Aguiló, J.; De-la Camara, C.; Bailón, R.; Scilingo, E.P. Acute stress state classification based on electrodermal activity modeling. *IEEE Trans. Affect. Comput.* **2021**, doi:10.1109/TAFFC.2021.3055294. [[CrossRef](#)]
- Peternel, K.; Pogačnik, M.; Tavčar, R.; Kos, A. A presence-based context-aware chronic stress recognition system. *Sensors* **2012**, *12*, 15888–15906. [[CrossRef](#)]
- Zangróniz, R.; Martínez-Rodrigo, A.; Pastor, J.M.; López, M.T.; Fernández-Caballero, A. Electrodermal activity sensor for classification of calm/distress condition. *Sensors* **2017**, *17*, 2324. [[CrossRef](#)]
- Liu, Y.; Du, S. Psychological stress level detection based on electrodermal activity. *Behav. Brain Res.* **2018**, *341*, 50–53. [[CrossRef](#)] [[PubMed](#)]
- Arza, A.; Garzón-Rey, J.M.; Lázaro, J.; Gil, E.; Lopez-Anton, R.; de la Camara, C.; Laguna, P.; Bailón, R.; Aguiló, J. Measuring acute stress response through physiological signals: towards a quantitative assessment of stress. *Med. Biol. Eng. Comput.* **2019**, *57*, 271–287. [[CrossRef](#)] [[PubMed](#)]
- Cerutti, S.; Hoyer, D.; Voss, A. Multiscale, multiorgan and multivariate complexity analyses of cardiovascular regulation. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2009**, *367*, 1337–1358. [[CrossRef](#)] [[PubMed](#)]
- Murthy, J.N.; Van Jaarsveld, J.; Fei, J.; Pavlidis, I.; Harrykissoon, R.I.; Lucke, J.F.; Faiz, S.; Castriotta, R.J. Thermal infrared imaging: A novel method to monitor airflow during polysomnography. *Sleep* **2009**, *32*, 1521–1527. [[CrossRef](#)]
- Garbey, M.; Sun, N.; Merla, A.; Pavlidis, I. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *IEEE Trans. Biomed. Eng.* **2007**, *54*, 1418–1426. [[CrossRef](#)]
- Merla, A.; Di Donato, L.; Romani, G.; Proietti, M.; Salsano, F. Comparison of thermal infrared and laser doppler imaging in the assessment of cutaneous tissue perfusion in scleroderma patients and healthy controls. *Int. J. Immunopathol. Pharmacol.* **2008**, *21*, 679–686. [[CrossRef](#)]
- Di Giacinto, A.; Brunetti, M.; Sepede, G.; Ferretti, A.; Merla, A. Thermal signature of fear conditioning in mild post traumatic stress disorder. *Neuroscience* **2014**, *266*, 216–223. [[CrossRef](#)]
- Engert, V.; Merla, A.; Grant, J.A.; Cardone, D.; Tusche, A.; Singer, T. Exploring the use of thermal infrared imaging in human stress research. *PLoS ONE* **2014**, *9*, e90782. [[CrossRef](#)]
- Genno, H.; Ishikawa, K.; Kanbara, O.; Kikumoto, M.; Fujiwara, Y.; Suzuki, R.; Osumi, M. Using facial skin temperature to objectively evaluate sensations. *Int. J. Ind. Ergon.* **1997**, *19*, 161–171. [[CrossRef](#)]
- Or, C.K.; Duffy, V.G. Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement. *Occup. Ergon.* **2007**, *7*, 83–94. [[CrossRef](#)]

23. Salazar-López, E.; Domínguez, E.; Ramos, V.J.; De la Fuente, J.; Meins, A.; Iborra, O.; Gálvez, G.; Rodríguez-Artacho, M.; Gómez-Milán, E. The mental and subjective skin: Emotion, empathy, feelings and thermography. *Conscious. Cogn.* **2015**, *34*, 149–162. [[CrossRef](#)]
24. Filippini, C.; Spadolini, E.; Cardone, D.; Merla, A. Thermal Imaging Based Affective Computing for Educational Robot. In Proceedings of the Multidisciplinary Digital Publishing Institute Proceedings, Florence, Italy, 17–19 September 2019; Volume 27, p. 27.
25. Sonkusare, S.; Ahmedt-Aristizabal, D.; Aburn, M.J.; Nguyen, V.T.; Pang, T.; Frydman, S.; Denman, S.; Fookes, C.; Breakspear, M.; Guo, C.C. Detecting changes in facial temperature induced by a sudden auditory stimulus based on deep learning-assisted face tracking. *Sci. Rep.* **2019**, *9*, 4729. [[CrossRef](#)] [[PubMed](#)]
26. Goulart, C.; Valadao, C.; Delisle-Rodriguez, D.; Caldeira, E.; Bastos, T. Emotion analysis in children through facial emissivity of infrared thermal imaging. *PLoS ONE* **2019**, *14*, e0212928. [[CrossRef](#)] [[PubMed](#)]
27. Cho, Y.; Bianchi-Berthouze, N.; Oliveira, M.; Holloway, C.; Julier, S. Nose heat: Exploring stress-induced nasal thermal variability through mobile thermal imaging. In Proceedings of the 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), Cambridge, UK, 3–6 September 2019; pp. 566–572.
28. Yamakoshi, T.; Yamakoshi, K.I.; Tanaka, S.; Nogawa, M.; Shibata, M.; Sawada, Y.; Rolfe, P.; Hirose, Y. A preliminary study on driver's stress index using a new method based on differential skin temperature measurement. In Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007; pp. 722–725.
29. Puri, C.; Olson, L.; Pavlidis, I.; Levine, J.; Starren, J. StressCam: Non-contact measurement of users' emotional states through thermal imaging. In Proceedings of the CHI'05 Extended Abstracts on Human Factors in Computing Systems, Portland, OR, USA, 2–7 April 2005; pp. 1725–1728.
30. Gioia, F.; Pascali, M.A.; Greco, A.; Colantonio, S.; Scilingo, E.P. Discriminating Stress From Cognitive Load Using Contactless Thermal Imaging Devices. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Guadalajara, Mexico, 1–5 November 2021; pp. 608–611.
31. Cruz-Albarran, I.A.; Benitez-Rangel, J.P.; Osornio-Rios, R.A.; Morales-Hernandez, L.A. Human emotions detection based on a smart-thermal system of thermographic images. *Infrared Phys. Technol.* **2017**, *81*, 250–261. [[CrossRef](#)]
32. Resendiz-Ochoa, E.; Cruz-Albarran, I.A.; Garduño-Ramon, M.A.; Rodriguez-Medina, D.A.; Osornio-Rios, R.A.; Morales-Hernández, L.A. Novel expert system to study human stress based on thermographic images. *Expert Syst. Appl.* **2021**, *178*, 115024. [[CrossRef](#)]
33. Merla, A.; Romani, G.L. Functional infrared imaging in medicine: A quantitative diagnostic approach. In Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 224–227.
34. Stroop, J.R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **1935**, *18*, 643. [[CrossRef](#)]
35. Cohen, S.; Kamarck, T.; Mermelstein, R. Perceived stress scale. In *Measuring Stress: A Guide for Health and Social Scientists*; Oxford University Press: Oxford, UK, 1994; Volume 10, pp. 1–2.
36. Fernández-Cuevas, I.; Marins, J.C.B.; Lastras, J.A.; Carmona, P.M.G.; Cano, S.P.; García-Concepción, M.Á.; Sillero-Quintana, M. Classification of factors influencing the use of infrared thermography in humans: A review. *Infrared Phys. Technol.* **2015**, *71*, 28–55. [[CrossRef](#)]
37. Nirkin, Y.; Masi, I.; Tuan, A.T.; Hassner, T.; Medioni, G. On face segmentation, face swapping, and face perception. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; pp. 98–105.
38. Suhr, J.K. Kanade-lucas-tomasi (klt) feature tracker. In *Comput. Vis. (EEE6503)*; Yonsei University: Seoul, Korea, 2009; pp. 9–18.
39. Greco, A.; Valenza, G.; Lanata, A.; Scilingo, E.P.; Citi, L. cvxEDA: A convex optimization approach to electrodermal activity processing. *IEEE Trans. Biomed. Eng.* **2015**, *63*, 797–804. [[CrossRef](#)]
40. Posada-Quintero, H.F.; Florian, J.P.; Orjuela-Cañón, A.D.; Aljama-Corrales, T.; Charleston-Villalobos, S.; Chon, K.H. Power spectral density analysis of electrodermal activity for sympathetic function assessment. *Ann. Biomed. Eng.* **2016**, *44*, 3124–3135. [[CrossRef](#)]
41. Tarvainen, M.P.; Niskanen, J.P.; Lipponen, J.A.; Ranta-Aho, P.O.; Karjalainen, P.A. Kubios HRV—heart rate variability analysis software. *Comput. Methods Programs Biomed.* **2014**, *113*, 210–220. [[CrossRef](#)]
42. Pan, J.; Tompkins, W.J. A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **1985**, 230–236. doi:10.1109/TBME.1985.325532. [[CrossRef](#)] [[PubMed](#)]
43. Malik, M.; Camm, A.J. Heart rate variability. *Clin. Cardiol.* **1990**, *13*, 570–576. [[CrossRef](#)] [[PubMed](#)]
44. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 289–300. [[CrossRef](#)]
45. Yan, K.; Zhang, D. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sens. Actuators B Chem.* **2015**, *212*, 353–363. [[CrossRef](#)]
46. Tang, D.; Jin, W.; Qin, N.; Li, H. Feature selection and analysis of single lateral damper fault based on SVM-RFE with correlation bias reduction. In Proceedings of the 2016 35th Chinese Control Conference (CCC), Chengdu, China, 27–29 July 2016; pp. 3840–3845.

47. Saeys, Y.; Inza, I.; Larranaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **2007**, *23*, 2507–2517. [[CrossRef](#)] [[PubMed](#)]
48. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
49. Yu, L.; Liu, H. Feature selection for high-dimensional data: A fast correlation-based filter solution. In Proceedings of the 20th International Conference on Machine Learning (ICML-03), Washington, DC, USA, 21–24 August 2003; pp. 856–863.
50. Zhai, J.; Barreto, A. Stress detection in computer users based on digital signal processing of noninvasive physiological variables. In Proceedings of the 2006 International Conference of the IEEE Engineering In Medicine And Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 1355–1358.
51. Renaud, P.; Blondin, J.P. The stress of Stroop performance: Physiological and emotional responses to color—Word interference, task pacing, and pacing speed. *Int. J. Psychophysiol.* **1997**, *27*, 87–97. [[CrossRef](#)]
52. Kim, H.G.; Cheon, E.J.; Bai, D.S.; Lee, Y.H.; Koo, B.H. Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investig.* **2018**, *15*, 235. [[CrossRef](#)]
53. Őri, Z.; Monir, G.; Weiss, J.; Sayhouni, X.; Singer, D.H. Heart rate variability: Frequency domain analysis. *Cardiol. Clin.* **1992**, *10*, 499–533. [[CrossRef](#)]
54. Sloan, R.; Shapiro, P.; Bagiella, E.; Boni, S.; Paik, M.; Bigger Jr, J.; Steinman, R.; Gorman, J. Effect of mental stress throughout the day on cardiac autonomic control. *Biol. Psychol.* **1994**, *37*, 89–99. [[CrossRef](#)]
55. Choi, J.; Ahmed, B.; Gutierrez-Osuna, R. Development and evaluation of an ambulatory stress monitor based on wearable sensors. *IEEE Trans. Inf. Technol. Biomed.* **2011**, *16*, 279–286. [[CrossRef](#)] [[PubMed](#)]
56. Vlemincx, E.; Taelman, J.; De Peuter, S.; Van Diest, I.; Van Den Bergh, O. Sigh rate and respiratory variability during mental load and sustained attention. *Psychophysiology* **2011**, *48*, 117–120. [[CrossRef](#)] [[PubMed](#)]
57. Abdelrahman, Y.; Velloso, E.; Dingler, T.; Schmidt, A.; Vetere, F. Cognitive heat: Exploring the usage of thermal imaging to unobtrusively estimate cognitive load. In Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, New York, NY, USA, 11 September 2017; Volume 1, pp. 1–20.
58. Kirschbaum, C.; Pirke, K.M.; Hellhammer, D.H. The ‘Trier Social Stress Test’—A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* **1993**, *28*, 76–81. [[CrossRef](#)]
59. Panasiti, M.S.; Cardone, D.; Pavone, E.F.; Mancini, A.; Merla, A.; Aglioti, S.M. Thermal signatures of voluntary deception in ecological conditions. *Sci. Rep.* **2016**, *6*, 35174. [[CrossRef](#)]
60. Cho, Y.; Bianchi-Berthouze, N. Physiological and affective computing through thermal imaging: A survey. *arXiv* **2019**, arXiv:1908.10307.
61. Jatupaiboon, N.; Pan-Ngum, S.; Israsena, P. Subject-dependent and subject-independent emotion classification using unimodal and multimodal physiological signals. *J. Med. Imaging Health Inform.* **2015**, *5*, 1020–1027. [[CrossRef](#)]
62. Liu, H.; Zhou, M.; Liu, Q. An embedded feature selection method for imbalanced data classification. *IEEE/CAA J. Autom. Sin.* **2019**, *6*, 703–715. [[CrossRef](#)]
63. Lal, T.N.; Chapelle, O.; Weston, J.; Elisseeff, A. Embedded methods. In *Feature Extraction*; Springer: Berlin, Germany, 2006; pp. 137–165.
64. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [[CrossRef](#)]
65. Bagavathiappan, S.; Saravanan, T.; Philip, J.; Jayakumar, T.; Raj, B.; Karunanithi, R.; Panicker, T.; Korath, M.P.; Jagadeesan, K. Infrared thermal imaging for detection of peripheral vascular disorders. *J. Med. Phys.* **2009**, *34*, 43.
66. Abouelenien, M.; Burzo, M.; Mihalcea, R. Human acute stress detection via integration of physiological signals and thermal imaging. In Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments, Corfu Island, Greece, 29 June–1 July 2016; pp. 1–8.
67. Hong, K.; Hong, S. Real-time stress assessment using thermal imaging. *Vis. Comput.* **2016**, *32*, 1369–1377. [[CrossRef](#)]
68. Cho, Y.; Bianchi-Berthouze, N.; Julier, S.J. DeepBreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In Proceedings of the 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), San Antonio, TX, USA, 23–26 October 2017; pp. 456–463.
69. Nakanishi, R.; Imai-Matsumura, K. Facial skin temperature decreases in infants with joyful expression. *Infant Behav. Dev.* **2008**, *31*, 137–144. [[CrossRef](#)] [[PubMed](#)]
70. Marins, J.C.B.; Moreira, D.G.; Cano, S.P.; Quintana, M.S.; Soares, D.D.; de Andrade Fernandes, A.; da Silva, F.S.; Costa, C.M.A.; dos Santos Amorim, P.R. Time required to stabilize thermographic images at rest. *Infrared Phys. Technol.* **2014**, *65*, 30–35. [[CrossRef](#)]
71. Ammer, K.; Ring, E. Standard procedures for infrared imaging in medicine. In *Medical Devices and Systems*; CRC Press: Boca Raton, FL, USA, 2006; Volume 1.
72. Veltman, H.J.; Vos, W.W. Facial temperature as a measure of mental workload. In Proceedings of the 2005 International Symposium on Aviation Psychology, 2005; pp. 751–758. Available online: https://corescholar.libraries.wright.edu/isap_2005/138 (accessed on 29 December 2021).