ThermICA: Novel Approach for a Multivariate Analysis of Facial Thermal Responses

Federica Gioia*, Alberto Greco* *Senior Member, IEEE*, Alejandro Luis Callara *Member, IEEE*, Nicola Vanello, Enzo Pasquale Scilingo *Senior Member, IEEE*, and Luca Citi *Senior Member, IEEE*

***Abstract*— Objective: Infrared Thermography (IRT) has been used to monitor skin temperature variation in a contactless manner, in both clinical medicine and psychophysiology. Here, we introduce a new methodology to obtain information about autonomic correlates related to perspiration, peripheral vasomotility, and respiration from infrared recordings. Methods: Our approach involves a model-based decomposition of facial thermograms using Independent Component Analysis (ICA) and an ad-hoc preprocessing procedure. We tested our approach on 30 healthy volunteers whose psychophysiological state was stimulated as part of an experimental protocol. Results: Within-subject ICA analysis identified three independent components demonstrating correlations with the reference physiological signals. Moreover, a linear combination of independent components effectively predicted each physiological signal, achieving median correlations of 0.9 for electrodermal activity, 0.8 for respiration, and 0.73 for photoplethysmography peaks envelope. In addition, we performed a cross-validated inter-subject analysis, which allows to predict physiological signals from facial thermograms of unseen subjects. Conclusions/Significance: Our findings validate the efficacy of features extracted from both original and thermal-derived signals for differentiating experimental conditions. This outcome emphasizes the sensitivity and promise of our approach, advocating for expanded investigations into thermal imaging within biomedical signal analysis. It underscores its potential for enhancing objective assessments of emotional states.**

***Index Terms*— Blind Source Separation, Contactless Monitoring, Independent Component Analysis, Infrared Thermography, Psychophysiology, Skin Temperature.**

I. Introduction

INFRARED Thermography (IRT) is the technique that measures the infrared radiation emitted by objects with temperature above absolute zero. IRT allows for the detection of temperature variations and heat patterns of entire scenes, which are invisible to the human eye, enabling non-invasive and contactless temperature monitoring. In this context, the use of IRT in psychophysiology holds promise for monitoring mental states without physical contact to minimize the impact of electrodes on the mental state being investigated [1]–[4]. In particular, the observed variations in skin temperature, as detected by IRT, serve as a valuable indicator of Autonomic Nervous System (ANS) activity, reflecting the interplay of physiological factors such as subcutaneous blood perfusion, respiratory activity, metabolism, perspiration, and more. In this light, the thermal signal of the skin can be considered a multivariate signal, that carries information about several physiological processes.

Several previous studies have analyzed facial thermal images to extract features capable of characterizing the psychophysiological state [5]–[11]. Other investigations have attempted to isolate and retrieve physiological correlates from thermal recordings, such as perspiration, cutaneous blood perfusion, and respiration [12]–[16]. Specifically, the respiratory waveform has emerged as the most successfully retrieved physiological signal, achieved through various methods [15], [17], [18]. In most of the cases, these techniques have relied on manual or automatic segmentation and tracking of specific facial regions. However, one of the most promising techniques applied spatial Blind Source Separation (BSS) to isolate the nostrils and retrieve the associated temporal pattern, thereby capturing the respiratory signal [18]. Indeed, BSS decomposes a mixture of various source signals into their individual components, with no prior knowledge of the sources or the mixing process. This concept is applicable to a variety of fields, including audio, image, and biological signal processing. The Independent Component Analysis (ICA) is the most well-known BSS algorithm [19]. In a thermal video, each pixel records a variety of physiological and non-physiological (artefact) processes that affect the apparent temperature of the target region beneath it. Temporal ICA attempts to identify the individual temporally independent time courses from the observed mixtures and finds their associated spatial maps.

The aim of this study is to obtain information about the physiological phenomena that influence skin temperature, to gain a comprehensive understanding of the psychophysiological state of a subject at a distance. To this aim, we employed ICA to extract sources related to perspiration, peripheral vasomotility, and respiration from the thermal signals of individual facial pixels. ICA decomposition serves a dual purpose. Firstly, it separates the contributions of the physiological phenomena,

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”. The research leading to these results has received partial funding from the Italian Ministry of Education and Research (MIUR) in the framework of the FoReLab project and CrossLab project (Departments of Excellence).

Federica Gioia, Alberto Greco, Alejandro Luis Callara, Nicola Vanello and Enzo Pasquale Scilingo are with the Department of Information Engineering and the Research Center ”E. Piaggio” at the University of Pisa, Pisa, IT. Luca Citi is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester, UK. Corresponding author: Federica Gioia, e-mail: federicagioia94@gmail.com .

* = equally contributed authors

providing insights into various ANS mechanisms that would otherwise be challenging to discern through simple temperature evaluations due to potential overlap in their influence on skin temperature. Secondly, it generates spatial maps that enable the identification of patterns associated with each phenomenon, facilitating a non-invasive means of gaining physiological insights into thermal skin modulation across the face. This method represents an innovative approach in IRT in psychophysiology, allowing for the comprehensive analysis of facial thermograms, expanding beyond the standard region-specific approach [20], [21]. Indeed, this approach does not rely on the thermal signal extraction from a specific region of the face but it involves the whole face. Hence, this analysis is fully automatic, eliminating the need for region selection procedures. The main challenge of this approach is the need of motion correction. This is particularly important to make sure that each pixel produces a signal reflecting temperature changes rather than inherent movement in the face. Moreover, the ICA is sensitive to noise which can lead to lower accuracy of the extracted independent components or reduced separation quality. Here, we propose an ad-hoc preprocessing procedure to optimize the accuracy and reliability of the ICA model. Our approach was tested on the facial thermograms of 30 volunteers, using their physiological signals, i.e. electrodermal activity (EDA), photoplethysmography (PPG) and respiration (RESP), as reference for the independent sources obtained from their facial thermal recordings. We performed both inter- and intra- subject analyses to evaluate the physiological meaning of each independent source. The protocol was intended to activate the ANS, leading to alterations in the physiological signals. Furthermore, it was designed to induce distinct affective states to evaluate how well the independent sources obtained after the ICA can differentiate among them. See details in the next sections that are organized as follows: in Section II, we describe the experimental protocol and setup, the processing procedure of the physiological signals, the ICA model and the ad-hoc thermal processing procedure, the intra- and inter- subject analyses and the statistical analysis for feature comparison; then, we present experimental results and a discussion on them in Sections III and IV respectively.

II. Materials and Methods

A. Experimental Protocol and Setup

A cohort of 30 volunteers, with an age range of 26.6±3.6 years and comprising 20 females and 10 males, participated in this study. The experiments were approved by the "Bioethics Committee of the University of Pisa" (n.85 15/2019), and each volunteer signed the informed consent. The experimental timeline is shown in Fig. 1. The experiment involved alternating between rest sessions, Stroop test, and Emotional Elicitation Task (EET), each session lasting 5 minutes.

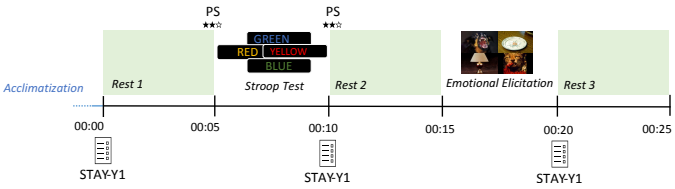

Fig. 1. Experimental timeline.

We designed a paced and computerized Stroop Test, intending to induce mental stress [22]. This is a well-known task that requires the resolution of stimuli with conflicting colour and semantic-meaning. Here, the subject had to press a button corresponding to the tint of the word presented in the middle of the screen every 2 seconds. Moreover, at the top of the screen, a counter kept track of the number of successes as a motivational stressor. Additionally, for each wrong/missing answer, the counter was reset while an acoustic buzzer alerted the subject. During the task, a ticking sound was being played in the background to mark the passage of time.

For the EET, we employed the International Affective Picture System (IAPS) public database, which is a consistent and effective resource for eliciting emotions [23]. It contains images rated by a large population in terms of valence, arousal, and dominance. The EET consisted of a set of images from the the IAPS, that had been chosen for their high level of arousal and negative valence, presented to the participant one after the other, for 5 seconds each.

Facial thermograms were acquired using a FLIR T640 thermal camera with a 24.6 mm lens, 640 × 480 pixels, Noise Equivalent Differential Temperature (NETD) < 0.04 mK @ +30°C and spectral range of 7.8–14 μm Long Wave InfraRed (LWIR). The sampling frequency was 5 Hz. The EDA and RESP were acquired using a BIOPAC MP150 system, at a sampling rate of 250 Hz. Specifically, EDA electrodes were placed on the distal phalanx of the index and ring fingers of the non-dominant hand. RESP was monitored with a piezoelectric chest belt. Finally, PPG was recorded using a Shimmer GSR + Unit sensor, with the optical pulse probe providing the signal from one finger of the non-dominant hand.

Regarding the thermal acquisition, we followed the checklist for the assessment of skin temperature detailed in [24]. In particular, the experiments were conducted in a controlled environment. The room atmospheric temperature and humidity ranged around 24–26 °C and 40–60%, respectively. Within the same experimental session, the temperature change was never over 1°C and humidity variation was a maximum of 5%. Moreover, all recordings took place away from direct heat and ventilation sources. Subjects were required to sit still and never move their head during the experiment, however, they did not have any physical constraints. They sat for at least 10 minutes before the acquisition started to complete a time of acclimatization in the testing environment for thermal regulation, balance with the testing environment, and stability of the participants' emotional states. The camera's emissivity was set to 0.98, the recommended value for a human body surface. Twenty minutes before the start of the recording, the camera was turned on to allow for sensor stabilization.

B. Autonomic Correlates Processing Procedure

In this section, we provide an overview of the physiological signals gathered in this study, namely EDA, RESP, and PPG, along with the processing algorithms employed.

EDA processing: The EDA signal records variations in skin conductance, which are directly affected by the amount of sweat produced by the eccrine glands. These glands are under the control of the sympathetic branch of the ANS, making the EDA signal a valuable tool for assessing SNS dynamics. The EDA signal can be divided into two components: tonic and phasic, each characterized by different time scales. The tonic component represents a slowly changing aspect of the signal with a spectrum below 0.05 Hz, and it carries information about the overall psychophysiological state of the subject. On the other hand, the phasic component reflects short-term and event-related responses. In this study, we employ the cvxEDA model to separate the EDA signal into these two components. The cvxEDA model is based on physiological principles and interprets the raw EDA signal as a linear combination of tonic and phasic elements, with the inclusion of white Gaussian noise that accounts for both model prediction errors and any measurement errors or artifacts. This model is founded on a robust framework that integrates Bayesian statistics, convex optimization, and sparsity, allowing for the decomposition of the EDA signal without the need for additional pre- or post-processing steps, as described in [25]. Of note, only the tonic component of the EDA signal was employed in this analysis, as it has comparable dynamics with the thermal signals. Specifically, the tonic component was downsampled to 5 Hz to match the frequency rate of the thermal recordings. Finally, we computed the mean of the tonic component, referred to as M_{Tonic}, averaged within non-overlapped 30-second-long time windows.

RESP processing: The RESP was used as a further index of sympathovagal activity. The typical frequency range for respiratory activity is 0.15 to 0.5 Hz, considering 9 to 30 cycles per minute. Therefore, we used a bandpass filter to isolate the respiratory activity from other sources in the signal and we downsampled it to 5Hz to match the frequency rate of the thermal recordings. Finally, we estimated the respiratory rate F_{RESP} as the maximum peak in the frequency spectrum.

PPG processing: PPG is an optical method to measure variations in blood volume in the microvascular bed of tissue. The PPG signal has a pulsatile component, called AC, superimposed onto a quasi-DC component. The AC waveform has its fundamental frequency around 1 Hz, depending on heart rate. Its increase is associated with the higher light attenuation due to the increase in microvascular blood volume with each heartbeat, also called "peripheral pulse". On the other hand, the quasi-DC component is related to the average blood volume under the tissue and it contains frequencies in the spectrum below 0.04Hz [26]. This DC component varies slowly due to respiration, vasomotor activity and vasoconstrictor waves. Therefore, the PPG signal provides information about the cardiovascular system and peripheral vascular activity [27]. Moreover, the evolution of the pulsatile amplitudes of the PPG signal reflects changes in the peripheral vasomotor activity [28], [29]. In this work, the PPG signal underwent a multi-step processing procedure to obtain a signal that carries only the amplitude information, which we will henceforth denote as PPG Peaks Envelope. First, a Butterworth highpass filter with a cutoff frequency of 0.04Hz was applied to eliminate the baseline. Subsequently, the PPG peaks were detected using the algorithm developed by Jesús Lázaro, relying on a low-pass-differentiator filter, as detailed in [30]. Finally, the detected peaks were linearly interpolated at 5 Hz, to align with the IR recordings frequency. From the obtained PPG Peaks Envelope signal, we estimated the mean which we refer to as M_{PPG}.

C. Independent Component Analysis Model of Thermal Recordings

The ICA model is a generative statistical model that describes how the observed data x_j, considered as random variables, are generated by a mixing process of the sources s_i, as follows:

$$x_j = a_{j,1}s_1 + a_{j,2}s_2 + ... + a_{j,P}s_P \tag{1}$$

where P is the number of sources. The ICA model of a thermal recording is described by:

$$X = AS \tag{2}$$

where thermal signals of length K from P pixels are arranged in matrix $X \in \mathbb{R}^{P\times K}$, $A \in \mathbb{R}^{P\times P}$ is the mixing matrix with elements a_{ij}, and $S \in \mathbb{R}^{P\times K}$ are the sources. In particular, the s variables are latent as they cannot be observed directly, and the A matrix is unknown.

The inverse of the A matrix is called the "unmixing matrix" and it is conventionally denoted by W, where $W = A^{-1}$. This provides an estimate $\tilde{S}$ of the sources S as:

$$\tilde{S} = WX \tag{3}$$

We gain a good approximation of the sources if the estimation of the unmixing matrix is accurate. For simplicity, this model ignores all noise components and any time delays in the recordings. However, additive noise can be present and, in this specific application, there is no time delay between the recordings, which are pixels of the same frame.

In order for ICA to work, it is necessary to make the assumption that source signals are "statistically independent" in addition to being uncorrelated. For thermal data, we consider the sources to be independent because we assume them to reflect different physiological processes. Moreover, the sources should not be Gaussian for the ICA to be possible. In fact, if they were Gaussian, their joint distribution would be symmetric, thus, it would have no information on the directions of the columns of the mixing matrix A, which could not be estimated. Here, ICA was performed using an implementation of the FastICA processing applications, based on Hyvarinen's fixed point algorithm maximizing non gaussianity [31].

Ideally, the ICA works well when the number of recordings is equal to the number of sources. In this case, even if we cannot say exactly how many sources are mixed to produce the observed thermal signals, we can expect them to be less than the number of recordings (i.e. the number of pixels). This

is called undercomplete ICA, and it requires a dimensionality reduction step in order to reduce the redundancy in the recordings. The most widely used dimensionality reduction method in signal processing applications is the Principal Component Analysis (PCA). PCA is a statistical technique that converts a set of possibly correlated variables into a set of linearly uncorrelated variables called principal components (PC), using an orthogonal transformation. PCA is often used to reduce the dimensionality of a dataset while preserving as much variability as possible [32]. In this application, PCA is used as a standard whitening process in ICA analysis and for dimensionality reduction at the same time. Whitening transforms the mixing matrix into a new orthogonal one, thus reducing the number of parameters to be estimated and the complexity of the ICA problem.

Thermal Recordings Preprocessing: The pipeline for the preprocessing of thermal video for ICA analysis is illustrated in Fig. 2.

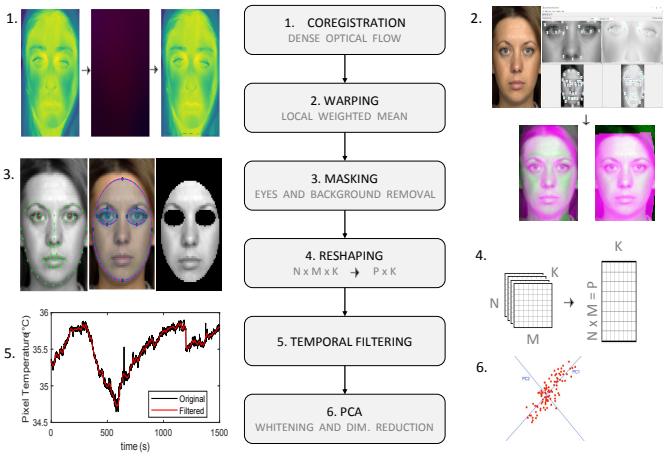

Fig. 2. Preprocessing pipeline for ICA.

First of all, the thermal frames were coregistered to remove motion. This is a crucial step to ensure that the ICA sees each pixel as a signal representing changes in temperature and is not misled by an underlying movement of the face. Therefore frames should be aligned at pixel level, so that corresponding pixels represent the same physiological features. Here, we used a dense optical flow algorithm with the adaption used in [18]. This method aligns the IR frames at the pixel level by estimating the nonlinear motion vector field between consecutive frames, effectively eliminating any movement of the subject.

To remove any physiognomic differences across participants, the first frame of each subject's IR video was then warped to match the anatomical features of a reference model. The RGB model was selected from the public database FACE [33], among the neutral faces. This step allows keeping the i-th pixel, ideally, in the same facial position for each participant throughout the inter-subject analysis. Since the IR frames for each subject had already been co-registered, eliminating possible movement artefacts, it was possible to calculate the transformation for each subject on the first frame and then apply it to all the frames of the same subject. In particular, a set of manually-selected fiducial points in the first frame of the IR video and the model image were used to estimate a transformation matrix which was applied to all the subsequent frames of the same video. This step relies on a Local Weighted Mean (LWM) transformation, which outperforms other geometrical transformations according to Cardone et al. [34]. All frames were spatially Gaussian smoothed to remove high-frequency noise. Moreover, we downsized the original IR frames to 60×80 pixels, effectively reducing its original size by a factor of 8, in order to lower the computational load. The total number of face pixels, or signals, after segmentation and eye masking was P = 2112.

A mask to remove the background and the eyes was applied to all frames before the analysis. The background is not of interest as we want to perform the analysis only on facial pixels, and the eyes are excluded to prevent blinks from contaminating the expected sources. In fact, blinks have a very strong impact on thermal pattern variations in the thermograms compared to the physiological temperature change of the skin. Specifically, the face segmentation mask was created by shaping an ellipse centred in the nose tip, with a minor axis length equal to the width of the face (being determined by the location of landmarks close to the ears) and a major axis length equal to twice the distance between the nose tip and the chin landmarks. Facial landmarks were identified in the RGB model image using the Yuval Nirkin algorithm [35]. Analogously, the ellipses to exclude the eyes were automatically determined based on the landmarks' location.

The processed videos were converted into matrices where each row corresponds to a pixel and each column to an observation in time. To clarify, if a frame has dimension $N \times M$ and the video has K frames, the output matrix will have $N \times M = P$ rows and K columns. In the inter-subject application, the matrices for all subjects were concatenated along the first dimension, resulting in a $P \times (K \cdot D)$ matrix, where D is the number of subjects.

The thermal signals were sampled at 5 Hz over a period of 25 minutes, hence, the number of observations was K = 7500. Each signal was filtered in time based on the specific target frequency range. More in detail, a lowpass filter with cut-off frequency equal to 0.05 Hz was applied to all signals when exploring the perspiration component, to match the frequency content of the tonic component of the EDA. Moreover, to isolate the respiration signal, we applied a bandpass filter with cutoff frequencies between 0.16 and 0.5 Hz. Whereas, no filter was applied to the signal when exploring the relation with the PPG Peaks Envelope signal.

Finally, whitening was performed together with dimensionality reduction by means of a PCA. As a result of the whitening process, the observed vector x is linearly transformed, its components are uncorrelated and their variances are equal to one. The PCA dimensionality reduction was set to use the first L eigenvectors, where L was equal to 10. These eigenvectors correspond to the directions capturing the highest variance within the high-dimensional data and projecting it onto a lower-dimensional subspace while keeping most of the information. The choice of L represents a trade-off between the complexity of the model and the achieved results.

Of note, when searching for a component associated with

the EDA tonic signal and the PPG Peaks Envelope, normalization by means of zscore was performed on the rows of the previously shaped $P \times K$ matrix of thermal signals. Indeed, considering the anticipated minor effects of processes such as perspiration and blood redistribution on the overall frame temperature compared to the impact of respiration, which leads to cold and warm airflow around the nostrils, this normalization enables the comparability of small variations to stronger factors.

D. Independent Components Evaluation

The thermal preprocessing led to a set of L Principal Components (PCs), uncorrelated between each other. After the application of the ICA model to these PCs, we obtained the ICs, temporally independent of each other. Our next step involved the selection of the components that correspond to each reference physiological signal. We performed both intra- and inter-subject analyses. In particular, we focused on retrieving the physiological components that are known to play a role in thermal modulation, while being also used in psychophysiology to infer the subject's mental state. Specifically, we used the tonic component of the EDA signal as a reference for the source related to the slow varying fluctuation of eccrine perspiration, the PPG Peaks Envelope for the peripheral vasomotor activity, and RESP for the respiration component.

Within-subject Analyses: Firstly, we performed the ICA for each subject, where the X matrix of the ICA model in (2) was composed of P rows, equal to the number of pixels of the face, and K columns, equal to the number of observations, as described in Section II-C. For each physiological signal (EDA, PPG, RESP), denoted as $y \in \mathbb{R}^{1\times K}$, we identified the j^{th} IC that exhibited the highest absolute value of the Spearman correlation coefficient, referred to as $\tilde{s}_{\max}$, which can be defined as:

$$\rho_{\tilde{s}_{\max} y} = \max_{1\leq j\leq L} \{|\rho_{\tilde{s}_j y}|\}. \tag{4}$$

The significance of the correlations was assessed using a non parametric permutation test with 1000 repetitions. In particular, to create the null distribution, the IC was split into 1-minute long blocks, which is reasonable for capturing meaningful temporal dynamics in physiological signals while introducing randomness to limit local dependencies, and they were randomly shuffled at each iteration. Analogously, we selected the j^{th} IC with the highest value of the cross-correlation coefficient, denoted as $\rho_{\tilde{s}_{\max} y}(\tau)$. This allows us to account for potential physiological delays between the reference and thermal signals, with τ representing the time displacement or lag between the respective time series. The $\rho_{\tilde{s}_{\max} y}(\tau)$ was evaluated for values of τ smaller than 15 seconds, as lags within this limit are likely to be physiological.

Furthermore, we computed the median values of both $\rho_{\tilde{s}\max y}$ and $\rho\tilde{s}_{\max} y(\tau)$, along with the Median Absolute Deviation (MAD), across all subjects.

At the subject level, we conducted another analysis that involves the use of all the ICs obtained after ICA. In particular, we fitted a linear model to explain the physiological signal y as a function of all the estimated ICs ($\tilde{s}_j$ with $j = 1, \ldots, L$):

$$y = c_0 + c_1\,\tilde{s}_1 + c_2\,\tilde{s}_2 + \cdots + c_L\,\tilde{s}_L \tag{5}$$

where c_j, with $j = 1, \ldots, L$, is the j^{th} coefficient, weighting of the j^{th} IC, and c_0 is the constant term within the model. The coefficient of determination R^2 was used as a metric to evaluate the goodness of fit of the model.

Inter-subject Analyses: In addition, we performed an inter-subject analysis to obtain a model that could generalize to new unseen subjects to extract the physiological signal from the processed IR videos. Importantly, this approach enables the extraction of components correlated with each physiological signal without relying on a reference, thereby making contactless analysis possible. To ensure the robustness of our model, we implemented Leave One Subject Out (LOSO) cross-validation of the pipeline including the PCA, ICA and linear regression model. At each iteration, we excluded one subject from the dataset, while using data from all other subjects as a training set ($D-1$ subjects). Specifically, we concatenated the X matrices corresponding to each subject along the first dimension, resulting in an extended data matrix, denoted as $X_{\text{train}} \in \mathbb{R}^{P\times R}$, where $R = K \cdot (D-1)$. Of note, this concatenation is only feasible due to the warping procedure, which registers all frames of all subjects to a common reference RGB model. This extended matrix X_{train} served as the input data for PCA, followed by subsequent ICA analysis. After obtaining the estimated sources $\tilde{s}_j$ through ICA, we used them to build a regression model, analogous to the model described in (5) for the within-subjects analysis, but fitted to $D-1$ subjects. The data from the excluded subject is used as the test set. The matrix $X_{\text{test}} \in \mathbb{R}^{P\times K}$ underwent a transformation based on the PCA and ICA outcomes obtained during the training phase. This transformation yields the predicted sources, denoted as $\hat{s}_j$. Subsequently, we compute the predicted physiological signal $\hat{y}$ as a linear combination of these $\hat{s}_j$ components:

$$\hat{y} = c_0 + c_1\,\hat{s}_1 + c_2\,\hat{s}_2 + \cdots + c_L\,\hat{s}_L \tag{6}$$

employing the coefficients c_j that were obtained on the training set. In order to quantify the relationship between the predicted $\hat{y}$ and the reference y, we performed the Spearman correlation coefficient between the two time series, $\rho_{y\hat{y}}$. Moreover, to assess the goodness of fit of our predictive model we computed the R^2, which measures the proportion of variance in the reference physiological signal y that is explained by our model. In addition, we computed the Root Mean Square Error (RMSE) which quantifies the average magnitude of the differences between the predicted and reference values, thus providing an understanding of the average accuracy of our predictions. For each physiological signal, the shared spatial maps across subjects were calculated using the whole dataset as a weighted sum of the map corresponding to each IC (6).

Moreover, to obtain the spatial map corresponding to each physiological signal $\hat{y}$, we obtained the coefficients c_j used in (6), by using the entire dataset (D subjects). These coefficients were then employed to weight the maps corresponding to each $\hat{s}_j$, resulting in the final representations. These maps can be

visually compared to the ones obtained from the raw thermal signals. Indeed, as a preliminary analysis, we examined the correlation between the raw thermal signals associated with individual pixels and each reference physiological signal [4]. Specifically, for each reference signal, a global correlation map was created by computing the average values of the correlation maps for all the participants, along with the standard deviation for each pixel. The correlation was evaluated by means of the Spearman correlation coefficient, denoted as ρ.

E. Features Comparison

Finally, we derived a representative feature from each couple of original and predicted physiological signal for the four distinct tasks. Specifically, we extracted the mean value from the original and predicted tonic component of the EDA signal, denoted as M_{Tonic} and $\widehat{M_{\text{Tonic}}}$, respectively. Similarly, we obtained the mean value of each original and predicted PPG Peaks Envelope signal, referred to as M_{PPG} and $\widehat{M_{\text{PPG}}}$, respectively. Additionally, from the original and predicted RESP signals, we computed the dominant frequency, denoted as F_{RESP} and $\widehat{F_{\text{RESP}}}$, respectively. We conducted a two-way Analysis of Variance (ANOVA) to assess differences between original and fitted features (a two-level variable, Feature), while simultaneously examining variations between conditions (a four-level variable, Task). In particular we tested three null hypothesis at the same time: I) there is no difference in group means at any level of the Task variable; II) there is no difference in group means of the Feature variable; III) the effect of one independent variable does not depend on the effect of the other independent variable. Finally, we performed the correlation and Bland-Altman plots to comprehensively assess the agreement between the original features and those extracted from the fitted signals.

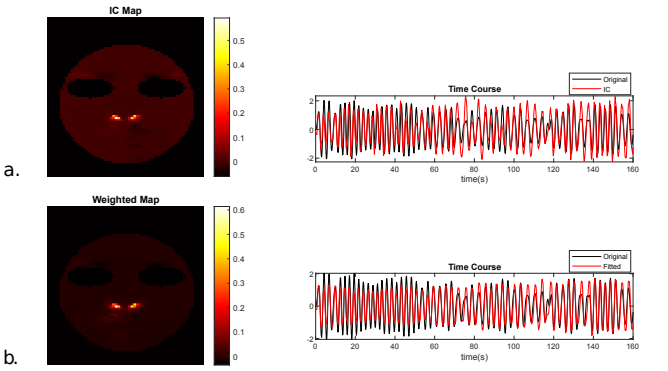

Fig. 4. Results of the ICA within subject between thermal and respiratory time series, for a sample subject.. a. Independent component (Right) and corresponding spatial map (Left); b. Estimated respiratory signal by a linear combination of independent components (Right) and corresponding spatial weighted map.

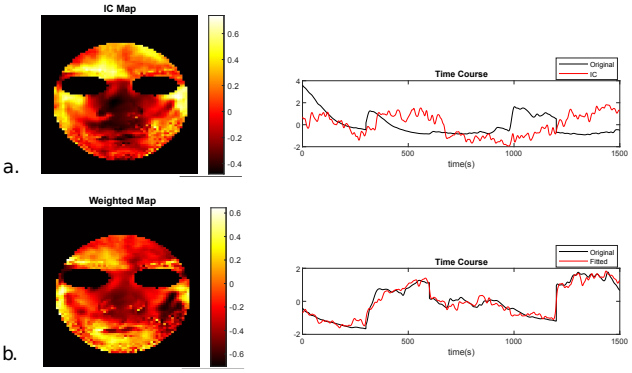

Fig. 3. Results of the ICA within subject between thermal and EDA tonic time series, for a sample subject. a. Independent component (Right) and corresponding spatial map (Left); b. Estimated EDA tonic component by a linear combination of independent components (Right) and corresponding spatial weighted map.

III. Results

Within-subject Analyses: In the context of the intra-subject analysis, the following results emerged. The results are reported as median ± MAD. For the EDA tonic signal, the $\rho_{\tilde{s}_{\max}y}$ across subjects was 0.49±0.1. The analogous $\rho_{\tilde{s}_{\max}y}(\tau)$ was 0.58±0.08, with lags of 7.4±4 seconds. Moreover, the coefficient of determination (R^2) of the linear regression model in (5) was 0.9±0.09 for the EDA tonic component. The IC $\tilde{s}_{\max}$ and the estimated $\hat{y}$ after the regression model along with the corresponding spatial and weighted map for a sample subject is shown in Fig. 3. While the signal was consistently recovered among all subjects, the maps exhibited variations from one subject to another.

For the RESP signal, the $\rho_{\tilde{s}_{\max}y}$ across subjects was 0.53±0.17, while the $\rho_{\tilde{s}_{\max}y}(\tau)$ was 0.80±0.14, with lags of 2±1 seconds. The R^2 of the linear regression model to explain the respiratory signal as a function of the $\tilde{s}_j$ was 0.73±0.17. Fig. 4 displays the IC $\tilde{s}_{\max}$, the estimated $\hat{y}$, and the related spatial maps for a sample subject. The signal was always accurately retrieved, and the nostrils were consistently depicted on the maps for all subjects. For the PPG Peaks Envelope signal, the $\rho_{\tilde{s}_{\max}y}$ across subjects was 0.42±0.08, while the $\rho_{\tilde{s}_{\max}y}(\tau)$ was 0.47±0.07, with lags of 6.6±2.3 seconds. The R^2 of the linear regression model to explain the respiratory signal as a function of the $\tilde{s}_j$ was 0.73±0.09. Fig. 5 displays the IC $\tilde{s}_{\max}$, the estimated $\hat{y}$, and the related spatial maps for a sample subject.

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3486628

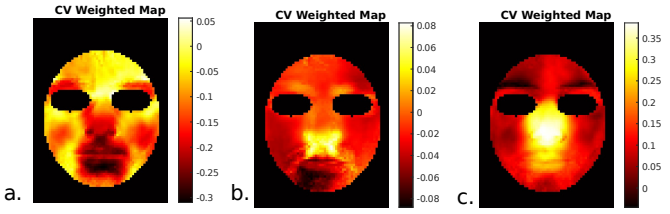

Fig. 6. Spatial Weighted Maps corresponding to each physiological signal $\hat{y}$, computed by using the entire dataset. a. EDA tonic component; b. RESP signal; PPG signal. The pixel values in the Spatial Weighted Maps represent the result of the linear combination of the values from the respective spatial maps associated to each individual IC.

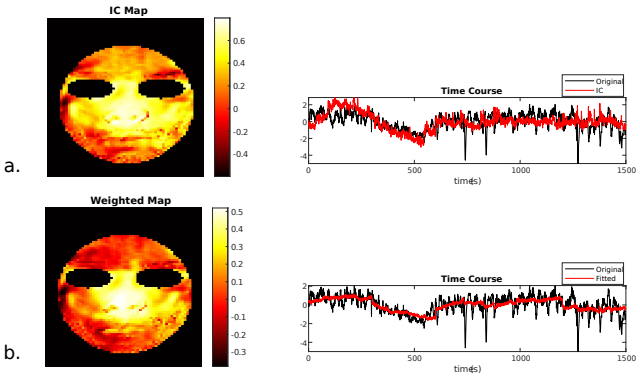

Fig. 5. Results of the ICA within subject between thermal and PPG Peaks Envelope time series, for a sample subject. a. Independent component (Right) and corresponding spatial map (Left); b. Estimated respiratory signal by a linear combination of independent components (Right) and corresponding spatial weighted map.

Inter-subject Analyses: Regarding the inter-subject analysis, the following findings were obtained for each physiological signal:

- EDA tonic: the median $\rho_{y\hat{y}}$ across the subjects was 0.45±0.19;
- RESP signal: the median $\rho_{y\hat{y}}$ across the subjects was 0.28±0.14;
- PPG Peaks Envelope signal: the median $\rho_{y\hat{y}}$ across the subjects was 0.4±0.15.

As additional metrics, we calculated the R^2 and RMSE for each physiological signal. For the EDA tonic signal, the median R^2 was 0.2 $\pm$ 0.24, with an RMSE of 0.9 $\pm$ 0.12. For the RESP signal, the median R^2 was 0 $\pm$ 0.08, with an RMSE of 1 $\pm$ 0.04. For the PPG Peaks Envelope signal, the median R^2 was 0.17 $\pm$ 0.15, with an RMSE of 0.91 $\pm$ 0.08.

The maps produced by the PCA, ICA, and Linear Regression pipeline using all the subjects are presented in Fig. 6.

The correlation maps between the thermal signals of the face and each physiological signal, along with the histograms of the standard errors per pixel, are shown in Fig. 7.

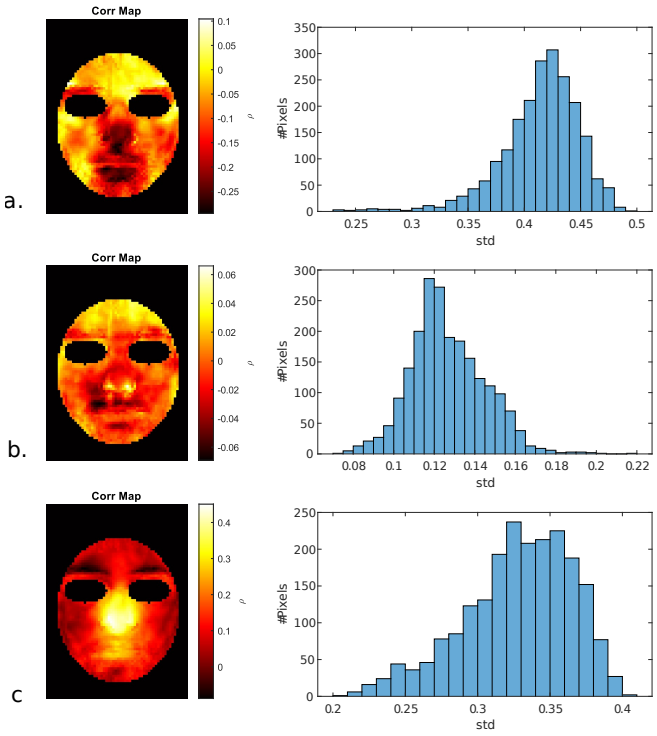

Fig. 7. Correlation Maps. Each pixel is the Spearman correlation coefficient between the thermal signal of that pixel and the physiological signal, and the histogram represents the number of pixels (#Pixels) per standard deviation value (std), for each physiological signal: a. EDA tonic; b. Respiratory signal RESP; c. PPG Peaks Envelope signal.

The tonic component of the EDA signal's correlation map reveals that the thermal signals of the pixels in the nose, maxilla, and chin areas are anti-correlated with the tonic signal, as shown by the presence of negative ρ values. The correlations with the respiratory signal are close to 0 across the entire face. The nostrils exhibit correlation values that are similar, in absolute value, to those observed on the forehead and upper lip. Positive correlations with the PPG Peaks Envelope signals are shown in the correlation map around the nose area.

A. Statistical Comparison

The results of the feature comparison between tasks are shown in Fig. 8. For all physiological features assessed, the comparison between original and fitted features did not reveal any statistically significant differences. Similarly, the interaction effect between the independent variables (Feature and Task) was not statistically significant. In contrast, across all scenarios, a significant Task effect was observed as the null hypothesis was rejected at the alpha level of 0.05.

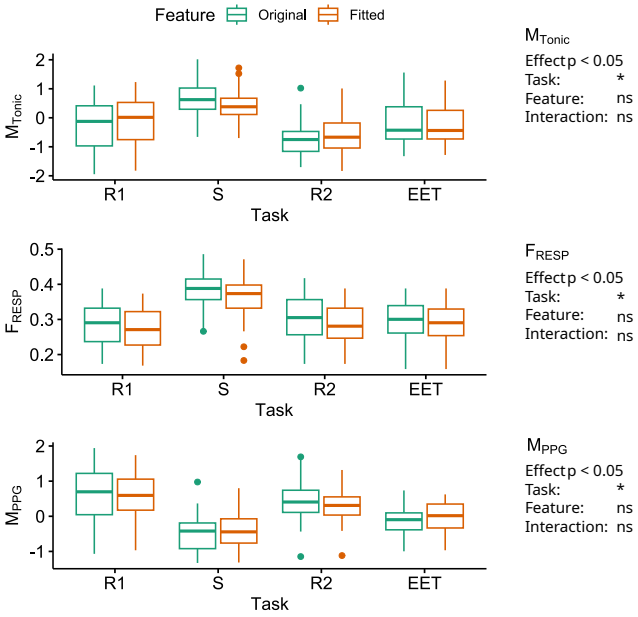

Fig. 8. Two-way ANOVA results for each physiological feature (M_{Tonic}, F_{RESP} and M_{PPG}); ns = non-significant; * = significant; R1 = First Rest; S = Stress; R2 = Second Rest; EET = Emotional Elicitation Task.

Particularly, the M_{Tonic} during R2 was significantly lower than in the other sessions ($p < .0001$ for all comparisons), whereas the M_{Tonic} during the stress session (S) was the highest ($p < .0001$ for all comparisons). R1 and EET showed intermediate levels that were significantly different from R2 ($p = .0078$ and $p = .0074$, respectively) and S ($p < .0001$ for both comparisons) sessions but not from each other ($p = 1.0000$). Concerning F_{RESP}, only S was statistically different from the other sessions, showing a significantly higher respiration rate ($p < .0001$ for all comparisons). Finally, for the feature M_{PPG}, no significant difference was found between R1 and R2 ($p = .3544$); however, the M_{PPG} during these resting phases was significantly higher than in the other sessions ($p < .0001$ for all comparisons). On the other hand, the S session showed a significantly lower M_{PPG} compared to EET ($p = .0006$), which consequently showed an intermediate level ($p < .0001$ for EET vs. R1 and $p = .0010$ for EET vs. R2).

B. Bland–Altman analysis results

Both M_{Tonic} and M_{PPG} features showed a strong linear relationship ($R^2 = 0.93$ and $R^2 = 0.92$ respectively) and no evidence of systematic bias (p>0.05) between the fitted and original values, supporting the robustness of our estimating method for both the EDA and PPG signals. In contrast, the F_{Resp} feature exhibited a moderate linear relationship ($R^2 = 0.72$) and a small but significant bias of 0.01 ($p < 0.001$). These results are visually summarized in the Bland-Altman and correlation plots for each feature in Figures 9.

IV. Discussions

Selecting a single IC has proven to be highly effective, consistently yielding correlation values exceeding 0.4 for all

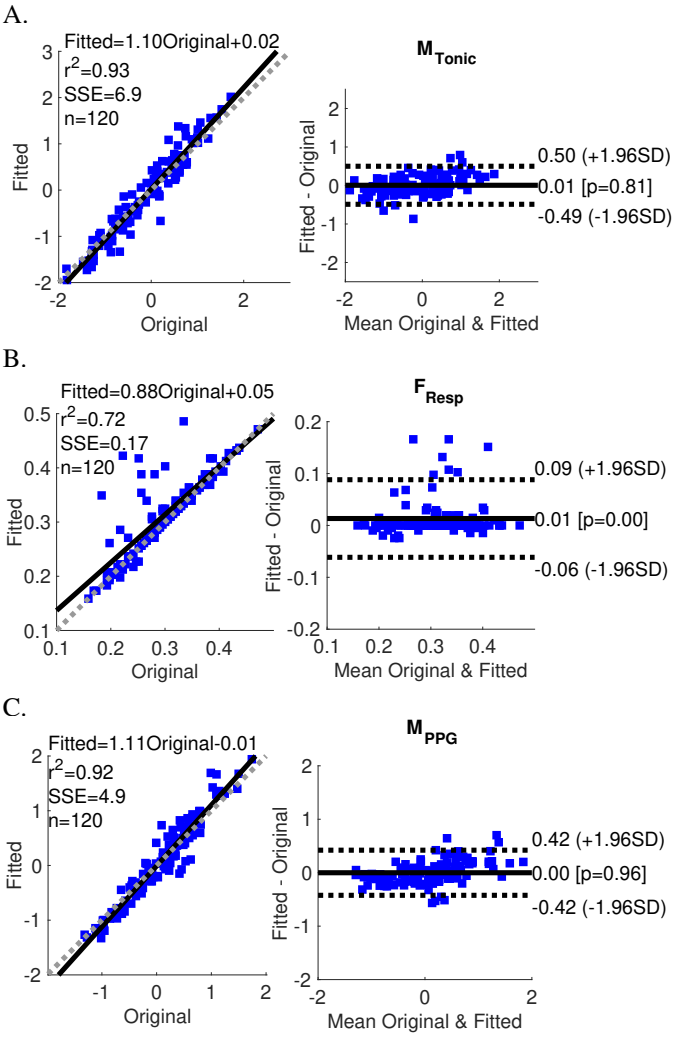

Fig. 9. Bland-Altman plot and correlation analysis for A) M_{Tonic}, B) F_{Resp}, C) M_{PPG} features.

the physiological signals under investigation. However, the within-subject analyses of the three reference signals revealed a noteworthy trend: the predictive performance significantly improved when employing a linear combination of all ICs to predict each physiological signal, in contrast to selecting only a single component. For instance, the correlation coefficients obtained between the predicted tonic component after the linear regression and the original signal reached a median of 0.9 with low variability across subjects. The correlation of one IC with the PPG peaks envelope signal, from a median of around 0.5, increases to 0.7 when all the components were linearly combined. These outcomes suggest that the ICA method employed might face challenges in effectively separating distinct physiological components, possibly identifying the same physiological phenomenon distributed across multiple sources. In practice, in situations where the unmixing model has more output ICs than truly independent source signals in the mixture, the excess components may appear as duplicates of unmixed ones close to independence. However, this is a common problem in ICA applications, and the ways to avoid it usually employ a stronger dimensionality reduction [36]. The

observed enhancement in predictive accuracy when combining multiple ICs underscores the complexity of the underlying physiological processes. In future work, as we anticipate a larger population sample, a model selection strategy will be employed to optimize the choice of the dimensionality reduction parameter, L. This strategy will involve assessing various values of L, using a cross-validation approach, to determine the dimensionality reduction that maximizes the effectiveness of our method.

In addition, it is worth noting that the correlations significantly increased in absolute value when examining a delayed version of the reference signals. This phenomenon is particularly important when assessing the relationship between the physiological response recorded with conventional systems in standard anatomical locations and the thermal response of the face. For instance, in the within-subject analyses, the EDA tonic component predominantly appeared in a single component, exhibiting a median correlation of around 0.49 consistently across subjects, as indicated by the low median absolute deviation. However, the correlation increased when considering a lag of 7.4 ± 4 seconds between the two signals. Temporal latency of cutaneous temperature change is a widely known phenomenon, but different delays have been documented by several studies [2]. More recently, another study reported a regionally specific decrease in temperature for up to 15 seconds and a short delay of approximately 2 seconds of thermal variation in comparison to EDA in response to an auditory stimulus, highlighting the need for deeper exploration [37]. Similarly, the correlation between the IC and the respiratory signal increased when considering a lag of around 2 seconds between the two, which is approximately the duration of a respiratory cycle.

On note, to evaluate and choose the component that was the most related to the referenced phenomenon, we used the absolute value of the correlation coefficient as one of the ambiguities of the ICA is the impossibility of retrieving the sign of the components, therefore one limitation of this method is the uncertainty of the correlation sign, and consequently the associated spatial maps could be as presented or the negative counterpart. In general, a very low MAD is associated with the median correlation values, meaning that the results are consistent across subjects.

As a preliminary analysis, we investigated the correlation maps between the thermal signals of each facial pixel and the reference signals, as in [4]. Each reference signal produced a unique mask. When examining the mask associated with the tonic component, a predominant pattern of negative coefficients emerged, particularly concentrated around the areas of the nose, upper lip, and chin. The temperature signal typically exhibits a negative correlation with skin conductance, which is consistent with other research in the field, because sweat production makes the conductance rise and the temperature decrease [4], [37]. The standard deviation associated with the map is high, indicating a wide variability among subjects. The map associated with the respiratory activity has low correlation values, ranging around zero. Even though the nostril regions displayed the highest correlation values, it is crucial to note that the generally low correlation values across the entire map preclude us from ascribing any meaningful interpretation to these results. Finally, the map associated with the peripheral vasomotor activity presents high correlations, close to 0.4, around the nose. This finding is in agreement with the results of our previous application which saw the nose as the most influenced region by the cardiovascular system during stress [38].

The spatial maps associated with the RESP component exhibit remarkable consistency across subjects, consistently pinpointing the nostrils or the region just below them. In contrast, the maps corresponding to perspiration and peripheral vasomotility display more diversity and are less straightforward to categorize visually. However, the cross-validation process allows us to derive distinct and unique maps for each physiological phenomenon that is shared across subjects. The incorporation of the warping procedure plays a pivotal role in this process. By aligning the anatomical features of all subjects, each pixel in the maps corresponds to the same anatomical location across time and subjects. Consequently, when we apply the maps generated using the training set to the test set, we can predict, on an unseen subject, physiologically meaningful time series data without the need for a reference signal. This marks a substantial advancement in the application of IRT in psychophysiology, as it enables us to generalize to new subjects without the requirement for individual calibration or electrode placement. Instead of relying on specific subject-related information, this method leverages the shared spatial distribution of each physiological phenomenon's impact on facial thermal variations. As a result, a set of established maps, generated from the training set, can be universally applied to new subjects for the acquisition of multiple physiological cues using just one contactless device. Therefore, by extracting information from the naturally occurring thermal variations on the face, we can access various physiological parameters related to the ANS activity. This approach not only simplifies the data collection process but also offers numerous practical applications in healthcare, where monitoring multiple physiological parameters is essential for patient care. The performance evaluation of the predicted versus the original data reveals promising results. Specifically, we obtain a correlation coefficient of 0.45 for the tonic component, 0.28 for the respiratory activity, and 0.4 for the envelope of the PPG peaks. These correlations underscore the effectiveness of our approach in predicting physiological signals patterns.

Regarding the determination coefficients, R^2, which measures the amount of variability of the dependent variable explained by the model and how well the model predicts the outcome, we observed lower values. In particular, the models explaining the EDA tonic and the PPG peaks envelope signals showed similar R^2, with a high subject variability. For the respiration, the R^2 was zero, meaning that the independent variables fail to explain any of the variation in the dependent variable. However, we noticed a scaling problem and potential phase shift between the predicted and original signal, which explains the low determination coefficient. Specifically, the model's inability to capture the amplitude variations of the respiratory signal contributed to this result, as well as possible phase shifts that may arise from differences in how the respira-

tory signals were measured. The position of the piezoelectric belt used to acquire the respiratory signal from the Biopac system can affect the phase, with belts positioned higher on the chest being more influenced by thoracic breathing, and those positioned lower capturing more of the abdominal component. Likewise, different facial and nose conformations among subjects may introduce variable delays in the thermal signal, as the thermal data reflect the cumulative warming effect of the airflow in the area beneath the nostrils. Nevertheless, we can accurately predict the frequency of the respiratory signal, which is considered a robust feature in psychophysiology, as demonstrated by the two-way ANOVA. This indicates that while the model may fail to capture the precise magnitude or timing of the respiration signal, it effectively identifies the periodicity and pattern of the respiratory cycles.

In addition, it is worthwhile noting that the map associated with the respiratory activity does not depict a well-defined area, as in the intra-subject maps, but highlights the lower part of the nose with two flows corresponding to the nostrils. Even in this case, this result may be due to the anatomical differences in the noses of different subjects. These differences could lead to variations in the thermal signals captured from the nostrils, depending on their visibility and the individual's facial structure. In fact, due to the conformation of their noses or the position of their faces, some subjects have less visible or completely hidden nostrils in the frames compared to others. For these subjects, the most responsive area will be the part immediately below the nose. Additionally, other facial features, such as the presence of a beard, can introduce variability in the thermal signal due to differential heat retention and reflection, which cannot be equalized using morphing procedures. Despite our population being heterogeneous, we were still able to obtain promising results. In fact, the correlation coefficient between the predicted and original signals were reasonably high. Indeed, the last analysis showed that the features obtained from the predicted signal were not significantly different from the features obtained from the original signals, as indicated by the comparison between features, original and fitted, in the two-way ANOVA. This outcome suggests that the ICA and regression modelling process effectively captured essential characteristics of the original data. In other words, our model successfully mimicked the features present in the original signals. At the same time, we observed a significant Task effect, meaning that these features varied significantly across the stimulation tasks as expected. Furthermore, the strong correlations and minimal biases showed by the Bland Altman analysis for the M_{Tonic} and M_{PPG} features indicate the effectiveness of our thermal imaging-based estimation method for these signals. On the other hand, the lower correlation and significant bias for F_{Resp} is consistent with previous limitations, suggesting a need for further refinement to enhance the robustness and accuracy of our method for this feature.

In the future, for real-case applications, a set of models, each adopting slightly different strategies (e.g., addressing beard or invisible nostrils), could be developed. For each new subject, the most suitable model from this diverse set could be selected based on the subject's distinctive morphology. This approach would overcome the limitations associated with relying on a single super-subject model. Further future works will focus on the integration of pulse rate variability (PRV) metrics to enhance the robustness of our physiological assessments from thermal imaging data. This will involve developing advanced algorithms for precise PPG raw signal estimation and peak detection directly from thermal videos. By achieving accurate peak detection, we can derive reliable PRV metrics, which are well-established indicators of autonomic nervous system function. Furthermore, the application of our method to higher-resolution thermograms will be explored. Specifically, reducing or eliminating the resizing step and finding a better balance between resolution and computational load could significantly enhance the performance of our analyses.

V. Conclusions

In conclusion, the application of temporal ICA for the direct prediction of physiological signals from thermal data is a remarkable breakthrough. Indeed, the proposed methodology involves a single, portable device for remote and unobtrusive monitoring of several physiological phenomena. As a result, this innovative approach not only contributes to the theoretical understanding of psychophysiological dynamics but also holds promise for real-world applications in fields such as healthcare, psychology, neurology, and human-computer interaction. The data that support the findings of this study are available from the corresponding author, F.G., upon reasonable request.

References

[1] Changjiang He, Mahdi Mahfouf, and Luis A Torres-Salomao. Thermal imaging for psychophysiological state detection in the human-machine interface (hmi) control system. In *Informatics in Control, Automation and Robotics: 15th International Conference, ICINCO 2018, Porto, Portugal, July 29-31, 2018, Revised Selected Papers 15*, pages 214–229. Springer, 2020.

[2] Stephanos Ioannou, Vittorio Gallese, and Arcangelo Merla. Thermal infrared imaging in psychophysiology: potentialities and limits. *Psychophysiology*, 51(10):951–963, 2014.

[3] Chiara Filippini, Edoardo Spadolini, Daniela Cardone, and Arcangelo Merla. Thermal imaging based affective computing for educational robot. In *Multidisciplinary Digital Publishing Institute Proceedings*, volume 27, page 27, 2019.

[4] Daniela Cardone, Francesco Cerritelli, Piero Chiacchiaretta, David Perpetuini, and Arcangelo Merla. Facial functional networks during resting state revealed by thermal infrared imaging. *Physical and Engineering Sciences in Medicine*, pages 1–16, 2023.

[5] Andrea Di Credico, David Perpetuini, Pascal Izzicupo, Giulia Gaggi, Daniela Cardone, Chiara Filippini, Arcangelo Merla, Barbara Ghinassi, and Angela Di Baldassarre. Estimation of heart rate variability parameters by machine learning approaches applied to facial infrared thermal imaging. *Frontiers in Cardiovascular Medicine*, 9:893374, 2022.

[6] David Perpetuini, Andrea Di Credico, Chiara Filippini, Pascal Izzicupo, Daniela Cardone, Piero Chiacchiaretta, Barbara Ghinassi, Angela Di Baldassarre, and Arcangelo Merla. Is it possible to estimate average heart rate from facial thermal imaging? *Engineering Proceedings*, 8(1):10, 2021.

[7] Vladimir Kosonogov, Lucas De Zorzi, Jacques Honore, Eduardo S Martínez-Velázquez, Jean-Louis Nandrino, José M Martinez-Selva, and Henrique Sequeira. Facial thermal variations: A new marker of emotional arousal. *PloS one*, 12(9):e0183592, 2017.

[8] Federica Gioia, Alberto Greco, Alejandro Luis Callara, and Enzo Pasquale Scilingo. Towards a contactless stress classification using thermal imaging. *Sensors*, 22(3):976, 2022.

[9] Gabriel A Wallace, Niten Singh, Elina Quiroga, and Nam T Tran. The use of smart phone thermal imaging for assessment of peripheral perfusion in vascular patients. *Annals of vascular surgery*, 47:157–161, 2018.

[10] K Mutlu, J Esquivelzeta Rabell, P Martin Del Olmo, and Sebastian Haesler. Ir thermography-based monitoring of respiration phase without image segmentation. *Journal of neuroscience methods*, 301:1–8, 2018.

[11] Youngjun Cho, Nadia Bianchi-Berthouze, and Simon J Julier. Deepbreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 456–463. IEEE, 2017.

[12] Alan T Krzywicki, Gary G Berntson, and Barbara L O'Kane. A non-contact technique for measuring eccrine sweat gland activity using passive thermal imaging. *International journal of psychophysiology*, 94(1):25–34, 2014.

[13] Dvijesh Shastri, Manos Papadakis, Panagiotis Tsiamyrtzis, Barbara Bass, and Ioannis Pavlidis. Perinasal imaging of physiological stress and its affective potential. *IEEE Transactions on Affective Computing*, 3(3):366–378, 2012.

[14] Ido Muller, Zehava Ovadia-Blechman, Noam Moyal, Noa Darchi, Oshrit Hoffer, Moshe Halak, and Neta Rabin. Combining thermal imaging and machine learning to noninvasively characterize palm perfusion during local blood pressure changes. *Biomedical Signal Processing and Control*, 92:106109, 2024.

[15] Meng-Han Hu, Guang-Tao Zhai, Duo Li, Ye-Zhao Fan, Xiao-Hui Chen, and Xiao-Kang Yang. Synergetic use of thermal and visible imaging techniques for contactless and unobtrusive breathing measurement. *Journal of biomedical optics*, 22(3):036006–036006, 2017.

[16] João Jorge, Mirae Harford, Mauricio Villarroel, Sitthichok Chaichulee, Shaun Davidson, Eoin Finnegan, Samuel H Clark, J Duncan Young, Peter J Watkinson, and Lionel Tarassenko. Non-contact assessment of peripheral artery haemodynamics using infrared video thermography. *IEEE Transactions on Biomedical Engineering*, 68(1):276–288, 2020.

[17] Gaetano Scebba, Giulia Da Poian, and Walter Karlen. Multispectral video fusion for non-contact monitoring of respiratory rate and apnea. *IEEE Transactions on Biomedical Engineering*, 68(1):350–359, 2020.

[18] Saurabh Sonkusare, Michael Breakspear, Tianji Pang, Vinh Thai Nguyen, Sascha Frydman, Christine Cong Guo, and Matthew J Aburn. Data-driven analysis of facial thermal responses and multimodal physiological consistency among subjects. *Scientific reports*, 11(1):1–12, 2021.

[19] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430, 2000.

[20] Ismael Fernández-Cuevas, Joao Carlos Bouzas Marins, Javier Arnáiz Lastras, Pedro María Gómez Carmona, Sergio Piñonosa Cano, Miguel Ángel García-Concepción, and Manuel Sillero-Quintana. Classification of factors influencing the use of infrared thermography in humans: A review. *Infrared Physics & Technology*, 71:28–55, 2015.

[21] Youngjun Cho and Nadia Bianchi-Berthouze. Physiological and affective computing through thermal imaging: A survey. *arXiv preprint arXiv:1908.10307*, 2019.

[22] J Ridley Stroop. Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6):643, 1935.

[23] Peter J Lang, Margaret M Bradley, Bruce N Cuthbert, et al. *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. NIMH, Center for the Study of Emotion & Attention Gainesville, FL, 2005.

[24] Danilo Gomes Moreira, Joseph T Costello, Ciro J Brito, Jakub G Adamczyk, Kurt Ammer, Aaron JE Bach, Carlos MA Costa, Clare Eglin, Alex A Fernandes, Ismael Fernández-Cuevas, et al. Thermographic imaging in sports and exercise medicine: A delphi study and consensus statement on the measurement of human skin temperature. *Journal of Thermal Biology*, 69:155–162, 2017.

[25] Alberto Greco, Gaetano Valenza, Antonio Lanata, Enzo Pasquale Scilingo, and Luca Citi. cvxeda: A convex optimization approach to electrodermal activity processing. *IEEE Transactions on Biomedical Engineering*, 63(4):797–804, 2015.

[26] Stefan Rasche, Robert Huhle, Erik Junghans, Marcelo Gama de Abreu, Yao Ling, Alexander Trumpp, and Sebastian Zaunseder. Association of remote imaging photoplethysmography and cutaneous perfusion in volunteers. *Scientific Reports*, 10(1):16464, 2020.

[27] John Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*, 28(3):R1, 2007.

[28] Frédéric Bousefsaf, Choubeila Maaoui, and Alain Pruski. Peripheral vasomotor activity assessment using a continuous wavelet analysis on webcam photoplethysmographic signals. *Bio-medical materials and engineering*, 27(5):527–538, 2016.

[29] Meir Nitzan, Anatoly Babchenko, Boris Khanokh, and David Landau. The variability of the photoplethysmographic signal-a potential method for the evaluation of the autonomic nervous system. *Physiological measurement*, 19(1):93, 1998.

[30] Jesús Lázaro, Eduardo Gil, José María Vergara, and Pablo Laguna. Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children. *IEEE journal of biomedical and health informatics*, 18(1):240–246, 2013.

[31] Aapo Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3):626–634, 1999.

[32] Rasmus Bro and Age K Smilde. Principal component analysis. *Analytical methods*, 6(9):2812–2831, 2014.

[33] Meredith Minear and Denise C Park. A lifespan database of adult facial stimuli. *Behavior research methods, instruments, & computers*, 36:630–633, 2004.

[34] Daniela Cardone, Edoardo Spadolini, David Perpetuini, Chiara Filippini, Antonio Maria Chiarelli, and Arcangelo Merla. Automated warping procedure for facial thermal imaging based on features identification in the visible domain. *Infrared Physics & Technology*, 112:103595, 2021.

[35] Yuval Nirkin, Iacopo Masi, Anh Tran Tuan, Tal Hassner, and Gerard Medioni. On face segmentation, face swapping, and face perception. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 98–105. IEEE, 2018.

[36] Hyeonseok Kim, Justin Luo, Shannon Chu, Cedric Cannard, Sven Hoffmann, and Makoto Miyakoshi. Ica's bug: How ghost ics emerge from effective rank deficiency caused by eeg electrode interpolation and incorrect re-referencing. *Frontiers in Signal Processing*, 3:1064138, 2023.

[37] Saurabh Sonkusare, David Ahmedt-Aristizabal, Matthew J Aburn, Vinh Thai Nguyen, Tianji Pang, Sascha Frydman, Simon Denman, Clinton Fookes, Michael Breakspear, and Christine C Guo. Detecting changes in facial temperature induced by a sudden auditory stimulus based on deep learning-assisted face tracking. *Scientific reports*, 9(1):1–11, 2019.

[38] Federica Gioia, Mimma Nardelli, Enzo Pasquale Scilingo, and Alberto Greco. Autonomic regulation of facial temperature during stress: A cross-mapping analysis. *Sensors*, 23(14):6403, 2023.