

Article

Intrinsic Motivational States Can Be Classified by Non-Contact Measurement of Autonomic Nervous System Activation and Facial Expressions

Sae Kawasaki ¹, Koichi Ashida ¹, Vinh-Tiep Nguyen ², Thanh Duc Ngo ², Duy-Dinh Le ² , Hirokazu Doi ^{3,4,5,*} and Norimichi Tsumura ^{1,6}

¹ Graduate School of Advanced Integration Science, Chiba University, Chiba 263-8522, Japan

² Faculty of Computer Science, University of Information Technology, Vietnam National University, Ho Chi Minh City 71308, Vietnam

³ Department of Information and Management Systems Engineering, Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka 940-2188, Japan

⁴ Data Science and AI Innovation Research Promotion Center, Shiga University, 1-1-1 Baba, Hikone 522-8522, Japan

⁵ School of Science and Engineering, Kokushikan University, 4-28-1 Setagaya, Tokyo 154-8515, Japan

⁶ Hiroshima University Hospital, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan

* Correspondence: doidoih@vos.nagaokaut.ac.jp; Tel.: +81-258-46-9322

Abstract: Motivation is a primary driver of goal-directed behavior. Therefore, the development of cost-effective and easily applicable systems to objectively quantify motivational states is needed. To achieve our goal, this study investigated the feasibility of classifying high- and low-motivation states by machine learning based on a diversity of features obtained by non-contact measurement of physiological responses and facial expression analysis. A random forest classifier with feature selection yielded modest success in the classification of high- and low-motivation states. Further analysis linked high-motivation states to the indices of autonomic nervous system activation reflective of reduced sympathetic activation and stronger, more intense expressions of happiness. The performance of motivational state classification systems should be further improved by incorporating different varieties of non-contact measurements.

Keywords: motivation; non-contact measurement; pulse wave; facial expression; action uni; machine learning



Citation: Kawasaki, S.; Ashida, K.; Nguyen, V.-T.; Ngo, T.D.; Le, D.-D.; Doi, H.; Tsumura, N. Intrinsic Motivational States Can Be Classified by Non-Contact Measurement of Autonomic Nervous System Activation and Facial Expressions. *Appl. Sci.* **2024**, *14*, 6697. <https://doi.org/10.3390/app14156697>

Academic Editor: Douglas O'Shaughnessy

Received: 12 June 2024

Revised: 25 July 2024

Accepted: 26 July 2024

Published: 31 July 2024



Copyright: © 2024 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/>).

1. Introduction

Motivation is the primary driver of goal-directed behavior, and highly motivated individuals can make continuous efforts to achieve their final goals. In the psychological literature, motivation has been grouped into two classes depending on the type of motivator [1]. One can be motivated to take certain actions guided by external, often tangible, reinforcers, such as money and food. Motivation induced by reward is termed “external motivation” [1]. By contrast, humans and other species can be committed to certain activities owing to the enjoyment and interest in the activity itself, regardless of whether the activity leads to the attainment of an external reward. This type of motivation is termed “intrinsic motivation” [1]. Since a pioneering study on intrinsic motivation in rhesus monkeys [2], several studies have been conducted on human participants. These studies have consistently linked high levels of intrinsic motivation to performance improvement and expertise in many domains of mental and physical activity [2,3].

An overarching goal of educational engineering is to develop learning-assistive technology to help people learn novel materials efficiently. Considering the close link between

students' intrinsic motivation and academic achievement [4,5], it is indispensable to develop an educational system that sustains students' intrinsic motivation for learning.

Importantly, there are great individual differences in learning styles and the mastery of academic subjects. Thus, educational materials should be personalized for each learner to enhance their intrinsic motivation to learn. One promising solution to achieve this is real-time monitoring of intrinsic motivation. By objectively evaluating temporal fluctuations in intrinsic motivation, the presentation schedule and type of educational materials can be adaptively determined to sustain an appropriate level of motivation in learners, thereby increasing learning efficiency and productivity.

There are several established ways to ascertain whether a person is engaged in an ongoing activity out of intrinsic motivation or not. Activities out of intrinsic motivation are often accompanied by positive feelings of curiosity, enjoyment and interest and are more likely to be deemed as "fun" rather than "work" [3]. Based on these observations, researchers have often utilized self-reports of subjective feeling during an ongoing activity or participant categorization of the activity as "fun" or "work" as a surrogate measure of intrinsic motivation [6,7]. As a more objective measure of intrinsic motivation, some researchers relied on the time one spends on the activity of interest without any tangible reward [8]. In addition to the measurement of intrinsic motivation, experimental methodology to manipulate the state of intrinsic motivation is valuable in investigations into the mechanisms of intrinsic motivation. However, many previous studies on intrinsic motivation relied on surveys or field experiments [3]. Thus, there are few alternatives to experimentally induce states with high and low intrinsic motivation. Today, one of the most widely accepted methods to experimentally manipulate the state of intrinsic motivation is the stopwatch (SW)/watch-stop (WS) paradigm [9] first introduced in a seminal functional magnetic resonance imaging (fMRI) study investigating the neural mechanisms of the undermining effect [10], i.e., the detrimental effect of external rewards on intrinsic motivation [8]. In the stopwatch (SW) task in this paradigm, participants were instructed to stop a stopwatch at a predefined time by mentally tracking the time elapsed. This game-like task was enjoyable and induced a high level of intrinsic motivation. In the watch-stop (WS) task, the stopwatch stopped automatically, and the participants were instructed to respond by pressing a button as soon as the stopwatch stopped counting. Compared with the SW task, the WS task was less engaging and even boring, and it induced low levels of intrinsic motivation. Since its introduction, the SW/WS paradigm tasks have been successfully used in several psychophysiological studies on intrinsic motivation [11–13].

Relatively little attention had been paid to the method for detecting motivational states or the "diagnosis" of problems in learning motivation [14], as pointed out in [15]. Though the number is still limited, several previous studies proposed systems for estimating motivational states from objectively quantifiable measures of behavioral and physiological signals. For example, Organero et al. analyzed behavioral data on an e-learning platform and succeeded in predicting intrinsic motivation by analyzing student behavioral data [16]. Of particular note, Chattopadhyay et al. proposed a method for detecting intrinsic motivation using electroencephalography (EEG) signals [17]. Specifically, they detected motivational states with 88% accuracy based on EEG data with the help of a Residual-in-Residual Convolutional Neural Network that enables feature learning with relatively small-sized data.

Thus, motivation state prediction based on physiological data reflective of nervous system activation is promising. However, when a system for physiological measurement is deployed in actual educational settings, the need for electrode attachment may interfere with and hinder the learning process. Thus, it is desirable to develop a system for motivation estimation based on psychophysiological signals that can be acquired without the attachment of cumbersome sensors.

In this study, we aimed to develop a method for estimating motivational states based on non-contact measurements of physiological and behavioral (facial expression) information from images of faces taken from videos. To achieve this, we developed a classifier to

dissociate the states of high and low intrinsic motivation induced by the SW/WS paradigm tasks based on indices of autonomic nervous system activation and facial expressions extracted from videos featuring faces. As stated above, previous studies have indicated that a state of high intrinsic motivation is often accompanied by positive emotion [6,7]. Thus, one promising approach to classify states of intrinsic motivation is to quantify emotional states during an activity. In a dimensional model of emotion [18], emotional states, defined as emotional arousal and valence, can be mapped in a two-dimensional space. Many psychophysiological studies have linked a state of arousal to sympathetic nervous system activation [19–21]. In a state of high arousal, the sympathetic nervous system is activated relative to the parasympathetic nervous system [19–21]. This change in balance between sympathetic and parasympathetic nervous system activation leads to phasic change in cardiac and pulse wave responses in the periphery [19–21]. We have formerly shown that pulse wave response can be estimated from the fluctuation in facial skin coloration [22]. On the basis of these, we decided to adopt a non-contact measurement of pulse wave responses to quantify the arousal of participants. One's facial expression conveys varying information on one's emotional state [21], including the intensity of positive and negative valence. Thus, we decided to incorporate a non-contact estimation of facial expression intensity based on videos featuring faces of participants.

There are several alternative methods of cardiac response measurement such as electrocardiography (ECG) [23] and photoplethysmography (PPG) [24,25]. ECG is suitable for a variety of biomedical applications, including heart rate measurement, examination of heart rhythm, diagnosis of cardiac abnormalities and biometric identification [23]. PPG can be measured non-invasively and provides valuable information about cardiovascular, respiratory and nervous system functions. Relevant to this study, emotion recognition using EEG and PPG signals has attracted attention in recent years. For instance, Shahid et al. proposed a system that recognizes four emotional states (happiness, sadness, fear and disgust) based on the features extracted from ECG and PPG. The resultant model for emotion classification achieved an accuracy of 85.7% [26]. In another study, Lee et al. utilized a single PPG signal to classify valence and arousal [27]. They segmented the PPG signal obtained with a contact pulse wavemeter into short segments and submitted them into a 1-Dimensional Convolutional Neural Network for feature extraction and emotion classification. Their proposed model achieved 75.3% accuracy for valence and 76.2% accuracy for arousal, enabling emotion recognition in a short period of time. These studies demonstrate the usefulness of ECG and PPG in emotion recognition, but, to the best of our knowledge, no study to date has utilized peripheral measurements of the autonomic nervous system, such as ECG and PPG, for the prediction of motivational state.

ECG and PPG contain large amounts of information, but adhesion of electrodes or sensors onto skin surface is required to measure these signals. Therefore, it is impractical to measure physiological responses in the natural environment using these methods. By contrast, a non-contact measurement of pulse wave does not require the adhesion of sensors on a participant's skin and, as such, poses a relatively small burden on participants. Thus, our proposed system has the potential to be utilized in the quantification of psychophysiological responses under wide-ranging situations. For example, images of faces taken from videos can be collected remotely. Thus, motivation monitoring based on images of faces taken from videos potentially opens up a way to improve learning motivation in home schoolers.

2. Pulse Wave Feature Acquisition Method

In a stressed or highly aroused state, the sympathetic nervous system is activated, whereas the activation of the parasympathetic nervous system is inhibited [19,28]. It is well established that the activation of the autonomic nervous system, and hence the balance between sympathetic and parasympathetic nervous system activation, is reflected in the temporal pattern of cardiac pulsation. Consequently, researchers in many fields have relied on electroencephalography and pulse wave measurements to estimate a person's

internal and emotional state [19,20]. As an application of this principle, Okubo et al. estimated autonomic nervous system activation based on pulse-wave data and proposed a biofeedback system to maintain driver alertness [29].

Immediately after cardiac pulsation, fresh red arterial blood circulates throughout the body. Synchronously with this “pulse wave”, redness, the color of oxygenated hemoglobin, becomes slightly intensified in the surface of the skin. Thus, pulse waves can be extracted as a time series of hemoglobin component intensities in a non-contact manner by analyzing skin pigmentation captured in a video [28].

In this study, we estimated the intensity of the hemoglobin component from faces taken from videos captured by an RGB camera, as proposed by Kurita et al. [22], while the participants were engaged in an SW/WS task. We extracted the time and frequency domain features based on the pulse wave (as described in the “6. Calculation of Pulse Wave Features”).

3. Hemoglobin and Shade Separation from Faces in Images

Independent component analysis (ICA) is a multivariate analysis technique that estimates the original signal from a mixture of independent signals without any information about the signals themselves or their composition. Here, independence refers to the state in which variables are orthogonal to each other, and the distribution of the variables is invariant.

Tsumura et al. applied ICA to faces in images and separated the pigment components of the skin [30]. In this procedure, the RGB values of a face in an image are used as observation signals in the ICA to estimate the pigment density distribution of the face. Here, we established the relationship between the RGB values of a face in an image and skin pigment density.

Figure 1a shows the human skin model proposed by Tsumura et al. [30]. The human skin is composed of three major layers: the epidermis, dermis and subcutaneous tissue. Skin coloration is determined by various pigments, among which melanin and hemoglobin are the most important. Melanin is found in the epidermis, whereas hemoglobin is found in the dermis. Therefore, under the simplified assumption that the epidermis is composed of melanin and that the dermis is composed of hemoglobin, the spatial distributions of melanin and hemoglobin can be considered independent.

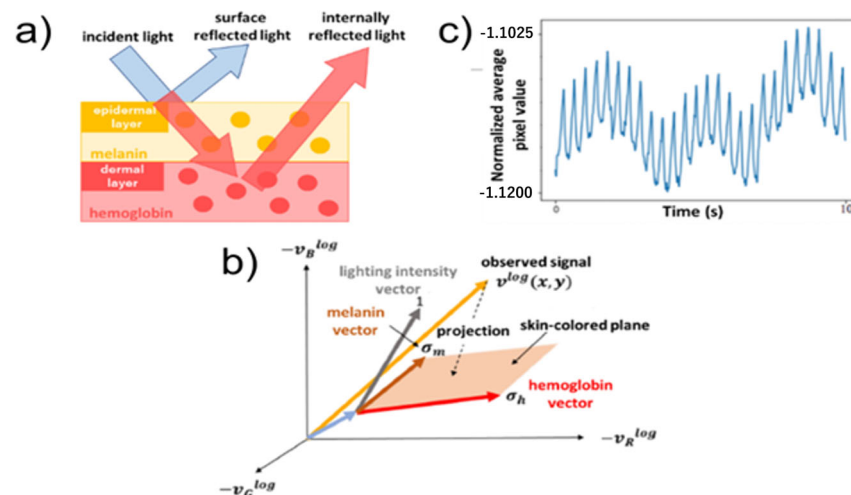


Figure 1. (a) Schematic representation of the skin model adopted in this study. (b) Observed signals and three independent signals [31]. (c) Example of bandpass-filtered pulse-wave signal.

Light incident on the skin is divided into surface-reflected light, which is reflected on the skin surface, and internally reflected light, which enters the skin and is repeatedly scattered before being emitted outside the skin. Surface-reflected light is not affected by skin pigmentation and expresses the color of the light source. By contrast, internally

reflected light represents the skin color because it is emitted after repeated absorption and scattering by pigments inside the skin.

Based on the optical skin model, the relationship between the pixel values and the observed signal (that is, the skin image) was extended to a mathematical model. The pixel values of the skin image can be defined as the sensor response of the camera in terms of the observed signal. The sensor response, v_i , of the camera is described by Equation (1),

$$v_i(x, y) = k \int L(x, y, \lambda) s_i(\lambda) d\lambda, \quad (1)$$

where i is an index identifying the color channel of the sensor; λ is the wavelength; $L(x, y, \lambda)$ is the spectral irradiance of the reflected light at position (x, y) on the skin surface; k is a constant indicating the camera gain; and s_i is the spectral sensitivity of the sensor. Applying the modified Lambert–Beer law to the spectral irradiance $L(x, y, \lambda)$ of the reflected light, Equation (2) can be derived from Equation (1).

$$v_i(x, y) = k \int e^{-\rho_m(x, y) \sigma_m(\lambda) l_e(\lambda) - \rho_h(x, y) \sigma_h(\lambda) l_d(\lambda)} E(x, y, \lambda) s_i(\lambda) d\lambda \quad (2)$$

where $E(x, y, \lambda)$ is the spectral irradiance of incident light at position (x, y) on the skin surface; ρ_m and ρ_h are the absorption cross section of melanin and hemoglobin, respectively; and l_e and l_d are the optical path lengths of light passing through the epidermis and dermis, respectively.

Since the spectral distribution of the skin is a smooth function along the wavelength axis and the spectral reflectance of the skin is highly correlated with the spectral sensitivity of the camera, the spectral sensitivity can be approximated as a narrow frequency band with $s_i(\lambda) \approx \delta(\lambda_i)$. Furthermore, assuming a sufficient distance between the illumination and the participant, the directional dependence of the illumination irradiance can be neglected. Assuming a single spectral characteristic of illumination, it is simplified to $E(x, y, \lambda) \approx p(x, y) \bar{E}(\lambda)$, where $p(x, y)$ is a variable indicating the change in shading owing to the skin shape. The above simplification transforms Equation (2) into Equation (3).

$$v_i(x, y) = k e^{-\rho_m(x, y) \sigma_m(\lambda) l_e(\lambda) - \rho_h(x, y) \sigma_h(\lambda) l_d(\lambda)} p(x, y) \bar{E}(\lambda) \quad (3)$$

Equation (3) is transformed into Equation (4) by performing a log transformation on the observed signal $v_i(x, y)$ and mapping it from a linear- to a log-scale space. The boldface in Equation (4) represents a vector.

$$\boldsymbol{v}^{log}(x, y) = -\rho_m(x, y) \sigma_m - \rho_h(x, y) \sigma_h + \boldsymbol{p}^{log}(x, y) \mathbf{1} + \boldsymbol{e}^{log} \quad (4)$$

where $\boldsymbol{v}^{log}$ denotes the log-transformed observed signal; (x, y) is the pixel position; $\rho_m(x, y)$ and $\rho_h(x, y)$ denote the density of melanin and hemoglobin, respectively; σ_m and σ_h denote the absorption cross-section of melanin and hemoglobin, respectively; $\boldsymbol{p}^{log}$ is the shading parameter due to skin shape; $\mathbf{1}$ is the intensity vector of shading; and $\boldsymbol{e}^{log}$ is the bias vector. To identify the independent signals, σ_m and σ_h , ICA was applied to the observed signals in a region where $\mathbf{1}$ and $\boldsymbol{e}^{log}$ were constant. Figure 1b shows the relationship between the observed and independent signals in the skin model.

4. Experiment

4.1. Participants

This study included 25 participants (5 females and 20 males) with a mean age of 24 years (standard deviation = 6.5). They participated in this study after giving written informed consent. The protocol of this study was approved by the ethical committee of the school of science and engineering in Kokushikan University (Approval No. R3-006).

4.2. Experimental Settings

Figure 2 shows an actual experimental scene. Participants were filmed while performing the assigned task on a laptop. Simultaneously, the pulse wave and EEG of the participants were measured. The results of the EEG data analysis will be reported elsewhere. An RGB camera (DFK33UX174, The Imaging Source, Bremen, Germany) was used to capture faces in images with a resolution of 640×480 px and a frame rate of 30 fps. The participants were photographed in a dark room with their faces fixed to a chin rest and as immobile as possible during the task. Two light-emitting diode (LED) lights (Viltrox L116T, Shenzhen Jueying Technology Co., Ltd., Shenzhen, China) were used as light sources, and the color temperature of all the lights was set to 5000 K. The two light sources were placed symmetrically around the participant.

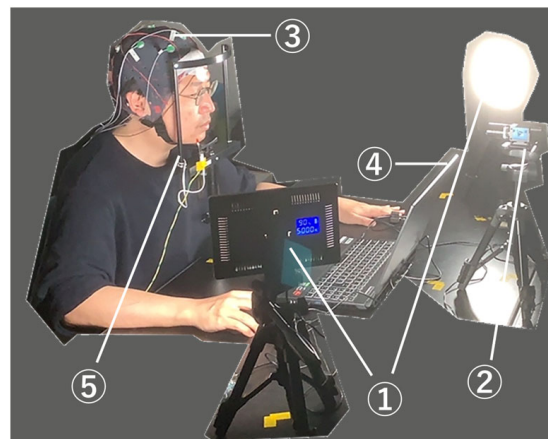


Figure 2. Experimental setting. 1: light source, 2: RGB camera, 3: EEG electrode cap, 4: laptop, 5: chin rest.

4.3. Task

In this study, we adapted the modified SW/WS tasks from a previous study [9]. In the SW task, the user tries to stop a stopwatch after exactly 5 s by mentally calculating the elapsed time. This task creates a state of increased motivation. If the display is stopped between 4.95 and 5.05 s, 1 point is added to the participant's score, and the total score is displayed at the top right corner of the screen. As previous studies have shown that people have the greatest sense of achievement for tasks of medium difficulty [32], this task was developed so that participants succeeded, on average, in half of the executions. A series of pilot experiments showed that this SW task is moderately difficult and intrinsically interesting for participants. By contrast, the WS task involves passively watching a stopwatch and clicking on a computer's mouse as soon as the stopwatch automatically stops after 5 s. There was no feedback of successes or failures in the WS condition. This simple task was intended to induce low motivation.

After each trial, the participant was prompted to provide a subjective rating of their motivational state. Nine white panels were presented horizontally on the screen. The leftmost panel represented the state of "Not at all motivated", and the right most panel represented "Highly motivated". The participants clicked the panel that they thought most appropriately reflected their feelings. After clicking a panel, the experiment proceeded to the next trial. We also collected subjective ratings of "the sense of achievement". The results of this rating will be reported elsewhere.

Both the WS and SW tasks included 30 trials each per participant. Recording in the WS and SW task lasted for approximately 6.5 min each. To avoid variations in the results due to the order in which the tasks were performed, the order of the WS and SW tasks was counterbalanced across the participants.

5. Data Set

A small red LED was placed next to the face of the participant and captured in the video together with participant's face. This LED was lit at the beginning and the end of each SW and WS task, respectively. Video frames submitted for further analysis were determined based on the timing of LED lighting. Specifically, the first frame was the frame just after the LED lighting at the beginning of a task, and the last frame was the frame just before the LED lighting at the end of a task.

Each task lasted for about 6.5 min. The number of frames used in the analysis varied from participant to participant because the duration of the task depended on the participant's response speed. The videos recorded were classified into high- and low-motivation states according to the subjective ratings (See "9. Machine Learning" below for the criteria for determining data with high- and low-motivation states). Table 1 shows the number of participants, the number of video data in each class and the number of frames.

Table 1. Description of the data set submitted for final analysis.

Number of participants	25	
Number of classes	2 (High- vs. low-motivation state)	
Number of video data in each class	High-motivation state	16
	Low-motivation state	34
Maximum number of frames	12,945 frames	
Minimum number of frames	10,564 frames	
Average number of frames	11,454 frames	
Standard deviation of the number of frames	556 frames	

6. Calculation of Pulse Wave Features

Pulse wave features were extracted from images of faces taken from videos by setting the region of interest (ROI), estimating pulse waves by hemoglobin and shade separation and extracting pulse wave features in the time and frequency domains.

First, a hemoglobin video image was obtained by applying dye-component separation to a facial video image captured using an RGB camera. Next, the participant's cheeks and nose were defined as the ROI, and the time series of the hemoglobin concentration components in this ROI were calculated as a pulse wave. The nose and cheeks were chosen because it is relatively easy to observe pulse wave changes in this facial region [33]. Hemoglobin values were averaged across pixels within a relatively large ROI for noise reduction.

Given that raw pulse wave signals contain considerable noise, trend removal [34] and bandpass filtering were applied as preprocessing for feature calculation. To improve the accuracy of peak detection, a bandwidth was set, and the peaks were detected. The bandwidth settings were obtained from Poh et al. [28]. Figure 1c shows an example of a signal preprocessed by trend removal and bandpass filtering, based on the time variation in the average pixel value in the ROI of a hemoglobin component image.

The peak-to-peak interval in the pulse wave can be considered an approximation to the RR interval in an ECG. Heart rate variability, which is the variation in the time intervals between successive R waves, is known to reflect the balance between sympathetic and parasympathetic nervous system activation. Therefore, a nonlinear analysis of the peak-to-peak intervals in the time and frequency domains could make it possible to assess the functioning of the autonomic nervous system.

Seven time domain features were obtained from estimated pulse wave signals. These included mean RR interval, standard deviation of RR intervals, mean heart rate, standard deviation of heart rate, Root Mean Square of Successive Differences (RMSSD), the number of successive RR intervals that differ by more than 50 ms (NN50) and the percentage of successive RR intervals that differ by more than 50 ms (pNN50) [35]. While the standard deviation of the RR interval reflects the degree of overall heart rate variability, the RMSSD reflects the degree of short-term heart rate variability, particularly variability due to parasympathetic activity. NN50 and pNN50 are parasympathetic indicators [35].

Time series of peak-to-peak intervals were Fourier-transformed, and ten frequency domain features, summarized in Table 2, were calculated to quantify parasympathetic and sympathetic activations. The converted low-frequency component (LF: 0.04–0.15 Hz) represents Mayer waves acting on both sympathetic and parasympathetic nerves, while the high-frequency component (HF: 0.15–0.4 Hz) shows a strong correlation with respiratory sinus arrhythmia, which is affected by respiration and parasympathetic action [34]. Therefore, LF/HF, LF/(LF + HF) and HF/(LF + HF) ratios can be used for sympathetic activity evaluation. Respiratory rate was estimated based on the frequency with the highest power spectrum density within HF range.

Table 2. Frequency domain features.

Respiratory Rate	Respiratory Rate per Minute
LF peak	Peak frequency below the upper limit of low-frequency band
HF peak	Peak frequency in the high-frequency band
ULF	Absolute power in the ultra-low-frequency band
VLF	Absolute power in the very-low-frequency band
LF	Absolute power in the low-frequency band
HF	Absolute power in the high-frequency band
LF(LF + HF)	Relative power in the low-frequency band
HF(LF + HF)	Relative power in the high-frequency band
LF/HF	LF and HF power ratio

In total, 17 different features were extracted from videos recorded of each participant. Some of these features had a high correlation (correlation coefficient ≥ 0.4). In the logistic regression analysis, collinearity among the features makes the analysis unstable. Similarly, in random forest, the presence of multiple highly correlated features makes the estimation of feature importance unstable and reduces the reliability of the analysis. Therefore, in this study, the pairwise correlations between features were reduced by manually removing some of the features according to the following criteria:

1. The ultra-low-frequency band (ULF) was removed because reliable data could only be obtained after 24 h of continuous measurement [35].
2. In the time domain, correlation coefficients between features exceeded 0.4 in almost all cases. So, first, we removed the features that were highly correlated with the mean value of the RR interval. This is because all other time domain features were computed based on this value. The standard deviation of heart rate was retained because of its low correlation with the RR interval mean.
3. Among the frequency domain features, those with correlation coefficients exceeding 0.4 with the RR interval were removed.
4. The correlation coefficient between HF peak and respiratory rate exceeded 0.4, so respiratory rate was removed. This is because respiratory rate is calculated from the HF peak.

Consequently, two time domain and two frequency domain features were retained: the mean peak-to-peak interval and standard deviation of heart rate as the time domain feature, and the LF peak and HF peak as frequency domain features. The correlation coefficient matrix between these features is shown in Figure 3.

The Variance Inflation Factor (VIF) was calculated; the VIF for the mean of the RR interval and the standard deviation of heart rate exceeded 20. Therefore, the standard deviation of heart rate was further removed according to rule 2 above. The VIFs of the retained features are summarized in Table 3.

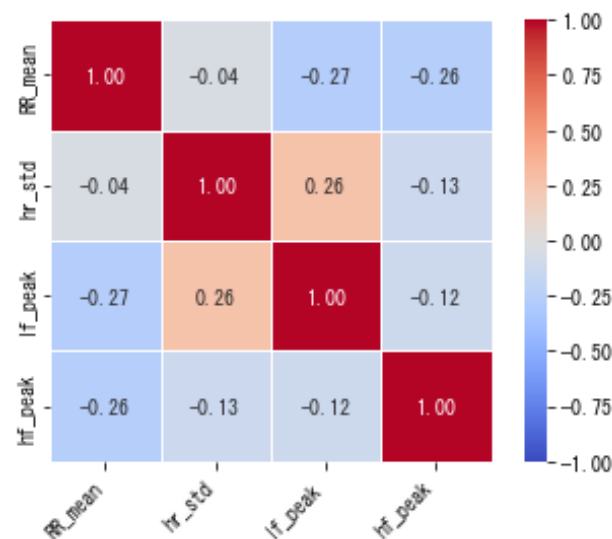


Figure 3. The correlational coefficient matrix of the features.

Table 3. VIFs of selected features.

Features	VIF
RR interval mean	11.5
LF peak	4.39
HF peak	9.58

7. Facial Expression Feature Calculations

Facial expressions were measured from videos featuring faces. In the “sign judgement” method of facial expression, facial muscle movements and actions are labelled as per definitions in, for example, the Facial Action Coding System (FACS) [36]. The FACS manual, revised in 2002, defines 27 action units (AUs), 14 head positions and movements, 9 eye positions and movements, 5 miscellaneous AUs, 9 action descriptors, 9 gross behaviors, and 5 visibility codes [37]. Each AU has five intensity ratings, ranging from A (minimum) to E (maximum). This coding scheme is assumed to provide the most objective assessment of facial movements.

In this study, OpenFace [38], an open-source automatic facial motion detection system, was used for facial expression estimation. The output of this toolkit includes the xy -axis motion of the head, facial tilt and degree of movement of the facial AU, as well as eye-gaze estimations, all of which are based on 68 feature points of the face. This software accepts videos, images, and live webcam feeds as input. Facial information was extracted frame by frame.

There are six facial expressions that are universal to all human beings: happiness, sadness, surprise, anger, fear and disgust [39]. In addition to these six basic facial expressions, the intensity of contempt, another universally observed facial expression, was evaluated. The intensities of these seven facial expressions were quantified as the sum of the AU intensity outputs using OpenFace. The correlation coefficients for sadness, fear and anger intensities exceeded 0.8, and only sadness, whose correlation coefficient was the lowest, was retained. Consequently, only the intensities of the facial expressions of happiness, sadness, surprise, disgust and contempt were used for further analysis. The AUs used to evaluate each facial expression are summarized in Table 4.

Table 4. Action unit definition for emotions [40].

Emotion	AU
Happiness	AU6 + AU12
Sadness	AU1 + AU4 + AU15
Disgust	AU9 + AU15 + AU16
Surprise	AU1 + AU2 + AU5 + AU26
Contempt	AU12 + AU14

8. Feature Selection

To avoid overfitting, feature selection was performed using two methods, Sequential Feature Selector (SFS) and Recursive Feature Elimination (RFE), implemented in scikit-learn (version 1.3.0) [41,42]. SFS is a method of increasing the number of features from a single-feature model to improve accuracy, while RFE is a method of removing less important features from a model of all features to improve accuracy.

As there were variations in the features selected between each run, feature selection runs were performed 100 times, and the frequency with which each feature was retained was counted. The most frequently retained features were used in the final model. We set the number of features to three, which yielded the highest F1 score.

9. Machine Learning

Selected pulse wave and facial expression features from the 25 participants were entered into machine learning algorithms as predictive features to classify high- and low-motivation states.

The motivational states were first classified into high- and low-motivation states based on subjective rating data based on the average scores for the 30 trials for each WS and SW task performed by each participant. Tasks with a subjective rating of motivation above 7 were defined as high-motivation states, whereas those with subjective ratings of motivation equal to or below 7 were defined as low-motivation states. In this study, the median rating was 5.3, representing the “neither motivated nor unmotivated” state. Therefore, the first quartile in the descending order of subjective ratings ≤ 7 was set as the threshold.

Next, three classification algorithms implemented in scikit-learn library, logistic regression [43], random forest classification [44] and support vector classification (SVC) [45] were tested for the classification of high- and low-motivation states. The parameters used in each model were tuned using a grid search algorithm. The initial parameters for grid search and the optimal parameters for each type of classifier and selection method are listed in Table 5. Learning and evaluation were performed 50 times by changing the random seed, and the best performance was adopted to avoid the local optimum.

Table 5. Initial parameters for grid search and the optimal parameters.

Model	Initial Parameters	Optimal Parameters for SFS	Optimal Parameters for RFE
Logistic regression	penalty = l2 C = 1.0	penalty = l2 C = 1.0	penalty = l2 C = 1.0
Random forest	n_estimators = 100 max_depth = None max_features = sqrt	n_estimators = 10 max_depth = 40 max_features = log2	n_estimators = 30 max_depth = 50 max_features = log2
SVC	kernel = rbf C = 1.0 gamma = scale	kernel = rbf C = 10 gamma = 1.0	kernel = rbf C = 1 gamma = 0.1

10. Results

The SFS- and RFE-selected features are listed in Table 6. As shown in this table, two features, RR interval mean and happiness, were consistently selected by both SFS and RFE.

Table 6. Features selected by SFS and RFE.

SFS	RFE
RR interval mean	RR intervals mean
Happiness	Happiness
LF peak	Contempt

Table 7 shows the results of the Brunner–Munzel test, which examines whether there is a statistically significant difference in the characteristics of each selected feature between the high- and low-motivation states. The results showed that the intensity of contempt was significantly smaller under low-motivation conditions (significance threshold of 0.05), while the differences in RR interval mean, LF peak, and happiness were not statistically different.

Table 7. Comparison of results between high- and low-motivation states.

Feature	Statistics	<i>p</i> -Value	Median (High)	Median (Low)
RR interval mean	−1.51	0.13	0.84	0.77
LF peak	1.32	0.19	0.037	0.044
Happiness	−1.39	0.17	1.66	0.99
Contempt	−2.38	0.024	1.31	0.56

To evaluate the classification performance, stratified K-split cross-validation (CV) was performed in triplicate to maintain the distribution of high- and low-motivation states in the training and test datasets.

Classification performance was evaluated using accuracy, specificity, area under the curve (AUC) in receiver operator characteristic (ROC) analysis and F1 score as performance indicators. Among these, the F1 score and AUC, which are robust against imbalanced data, are particularly important metrics.

The results of the stratified CV for SVC, logistic regression, and random forest classification using the features selected by SFS are summarized in Table 8. Additionally, the classification results obtained using the features selected by the RFE are summarized in Table 9.

Table 8. Classification performance metrics for classifiers using SFS.

	Threshold	Accuracy Rate	Specificity	AUC	F1
(a) Logistic Regression					
train	0.636	0.740	0.838	0.745	0.735
test	0.677	0.720	0.882	0.682	0.699
(b) Random Forest					
train	0.404	0.990	1.000	1.000	0.990
test	0.404	0.780	0.882	0.738	0.773
(c) SVC					
train	0.343	0.980	0.971	0.783	0.980
test	0.343	0.720	0.765	0.612	0.724

Table 9. Classification performance metrics for classifiers using RFE.

	Threshold	Accuracy Rate	Specificity	AUC	F1
(a) Logistic Regression					
train	0.586	0.690	0.750	0.733	0.693
test	0.667	0.720	0.853	0.686	0.708
(b) Random Forest					
train	0.404	1.000	1.000	1.000	1.000
test	0.434	0.740	0.853	0.760	0.732
(c) SVC					
train	0.394	0.760	0.853	0.517	0.755
test	0.343	0.640	0.618	0.671	0.652

As for the overall model classification performance, the random forest with features selected by RFE showed the best classification performance (AUC = 0.760). For the classification performance of the best discriminator, the random forest with features selected by SFS showed the best classification performance (F1 = 0.773).

As stated above, the mean RR interval and happiness were retained as features selected by both the SFS and RFE. Therefore, we repeated the machine learning algorithm using only these two features. The results are summarized in Table 10. The classification performance of classifiers using these two features outperformed the classification performance of random forest using the features selected by RFE and SFS (AUC = 0.788, F1 = 0.778).

Table 10. Classification performance metrics for classifiers using two features selected both by SFS and RFE.

	Threshold	Accuracy Rate	Specificity	AUC	F1
(a) Logistic Regression					
train	0.677	0.730	0.897	0.728	0.707
test	0.687	0.740	0.882	0.680	0.725
(b) Random Forest					
train	0.404	0.990	0.985	0.998	0.990
test	0.404	0.780	0.853	0.788	0.778
(c) SVC					
train	0.394	0.940	0.956	0.973	0.940
test	0.364	0.720	0.765	0.699	0.724

11. Discussion

This study investigated the feasibility of evaluating the state of intrinsic motivation by combining non-contact measurements of physiological responses and facial expressions. The proposed system achieved modest performance in the classification of high- and low-motivation states. Feature selection resulted in the consistent retention of the mean RR interval and the intensity of expressions of happiness. The heart rate variability indices reflect the activity of the autonomic nervous system [46–48], which in turn is deeply involved in the dopaminergic neural system [49], underlying the regulation of intrinsic motivational states [50,51]. In particular, all measures of autonomic activation are calculated from the mean RR interval, and the mean RR interval contains much information related to the autonomic nervous system. Therefore, it is not surprising that the mean RR interval was retained as predictive features by both SFS and RFE.

Statistical analysis between high- and low-motivation states showed a longer RR interval, although not statistically significant, in the high-motivation state than in the low-motivation state, indicating that reduced sympathetic activation constitutes one as-

pect of a highly motivational state. A traditional theory [52] has linked high levels of physiological arousal to stress, while a moderate level of physiological arousal has been linked to the optimal state of the so-called “flow state”, characterized by focused attention and full absorption in an ongoing activity. Thus, a numerically longer RR interval (hence, lower physiological arousal) during the high-motivation state raises the possibility that participants were more likely to have experienced the flow state during a high- rather than low-motivation state, while they tended to experience more stress during the low-motivation state. This interpretation dovetails the claim that the flow state experienced during an activity is associated with the intrinsic motivation to engage in and complete the activity [52,53].

After feature selection, the intensity of happiness was retained in both SFS and RFE, and contempt was retained in RFE. The classifier with the best performance (random forest using features selected by both SFS and RFE) used the intensity of happiness as a predictive feature. Though not statistically significant, a happy expression was numerically stronger in the high- rather than low-motivation state. This finding is consistent with previous studies linking positive emotions to intrinsic motivation [51]. In addition, it has been shown that smiling is an important signal of intrinsic motivation in terms of facial expressions [54]. A statistical comparison revealed higher levels of contempt in the high- rather than low-motivation state, which was not readily interpretable. At the AU level, both contempt and happy expressions include the constriction of the lip-corner pull (AU12) [39]. Thus, the effect of motivational state on contempt intensity might be a byproduct of AU 12 constriction. But this is just a speculative conjecture, and more information is required regarding the functional and expressive meanings of individual AUs to fully understand the present pattern of results.

In summary, this study revealed that the classification of motivational states can be achieved to a certain extent by combining a non-contact measurement of physiological responses and facial expression analysis. The classifiers mainly rely on the intensity of happy expressions and indices of autonomic nervous system activation, linking a high-motivation state to reduced sympathetic activation. However, reduced sympathetic activation and positive emotion may be associated in a non-specific manner with a positive psychological state. Consequently, an important task for future research is to establish one-to-one correspondence between objectively quantifiable physiological and highly motivational states by combining a wider range of multimodal psychophysiological information.

This study suffers from several limitations. First, we did not consider the effect of skin tone in the estimation of pulse wave from facial skin coloration. In a study by Ewa et al., heart rate was estimated using six different skin tones, with the darkest skin tone having the largest estimation error [55]. This study did not include participants with a particularly dark tone. Moreover, concern cannot be ruled out entirely that the present method failed to give an accurate estimate of the pulse wave in some of the participants. Second, with some relevance to the first limitation, this study included a small number of participants with a relatively homogenous background. Thus, the performance of the classifiers and its generalizability should be tested in a larger number of participants with more diverse attributes, e.g., ethnicity and age.

12. Conclusions

In this study, we used faces in images to classify the presence or absence of motivation and demonstrated the usefulness of the machine learning model and the features used. However, real-time estimation and continuous time series estimation are required for practical use. It is also necessary to evaluate motivation levels on a more minute scale rather than on a coarse binary classification. Therefore, in future studies, estimating data using regression models with time series data is required. This study is a kind of feasibility study showing the possibility of intrinsic motivation monitoring by non-contact measurement of facial information under laboratory settings. To establish a monitoring system that is useful in real-life settings, careful consideration should be given to at least two aspects.

First, this system utilizes very sensitive information of psychophysiological states extracted from data on faces. In addition, many recent studies have shown the vulnerability of machine learning applications to adversarial attacks [56]. Thus, security measures should be implemented in the system to keep the integrity of the system intact and protect against privacy intrusion. As for privacy breaches, a recent study [57] applied the technique of adversarial attack as a measure to protect an individual's private information from prying software. These techniques should be integrated into our system before testing it under real-life settings.

Second, learning is a long-lasting process; it sometimes takes several years to master a single subject. Thus, maintaining motivation throughout this long period is of critical importance for learners. This study tested the system's performance to estimate phasic states of intrinsic motivation. However, considering the long-lasting nature of the learning process, it is essential in future studies to examine if the longitudinal fluctuation of motivational states within an individual can be tracked by non-contact measurements of facial information.

Author Contributions: Conceptualization, H.D. and N.T.; methodology, H.D.; data collection, S.K., K.A. and H.D.; data analysis, S.K., K.A., H.D., V.-T.N., T.D.N. and D.-D.L.; writing—S.K. and H.D.; writing—review and editing, S.K., H.D. and N.T.; supervision, H.D. and N.T.; funding acquisition, H.D. and N.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Japan Science and Technology Agency (JST), grant number JPMJMI21D3.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the ethical committee of the school of science and engineering in Kokushikan University (Approval No. R3-006; Date of Approval: December 2021).

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

Data Availability Statement: Data will be made available by the corresponding author upon reasonable request.

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

References

1. Ryan, R.M.; Deci, E.L. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemp. Educ. Psychol.* **2000**, *25*, 54–67. [[CrossRef](#)] [[PubMed](#)]
2. Harlow, H.F.; Blazek, N.C.; McClearn, G.E. Manipulatory motivation in the infant rhesus monkey. *J. Comp. Physiol. Psychol.* **1956**, *49*, 444–448. [[CrossRef](#)] [[PubMed](#)]
3. Fishbach, A.; Woolley, K. The structure of intrinsic motivation. *Annu. Rev. Organ. Psychol. Organizational Behav.* **2022**, *9*, 339–363. [[CrossRef](#)]
4. Taylor, G.; Jungert, T.; Mageau, G.A.; Schattke, K.; Dedic, H.; Rosenfield, S.; Koestner, R. A self-determination theory approach to predicting school achievement over time: The unique role of intrinsic motivation. *Contemp. Educ. Psychol.* **2014**, *39*, 342–358. [[CrossRef](#)]
5. Howard, J.; Bureau, J.; Guay, F.; Chong, J.; Ryan, R. Student Motivation and Associated Outcomes: A Meta-Analysis From Self-Determination Theory. *Perspect. Psychol. Sci.* **2021**, *16*, 1300–1323. [[CrossRef](#)] [[PubMed](#)]
6. Amabile, T.M.; Hill, K.G.; Hennessey, B.A.; Tighe, E.M. The Work Preference Inventory: Assessing intrinsic and extrinsic motivational orientations. *J. Pers. Soc. Psychol.* **1994**, *66*, 5950–5967. [[CrossRef](#)] [[PubMed](#)]
7. Csikszentmihályi, M. *The Domain of Creativity. Theories of Creativity* 190–212; Sage: Thousand Oaks, CA, USA, 1990.
8. Lepper, M.R.; Greene, D.; Nisbett, R.E. Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *J. Pers. Soc. Psychol.* **1973**, *28*, 1129–1137. [[CrossRef](#)]
9. Murayama, K.; Matsumoto, M.; Izuma, K.; Matsumoto, K. Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20911–20916. [[CrossRef](#)] [[PubMed](#)]
10. Deci, E.L.; Koestner, R.; Ryan, R.M. The undermining effect is a reality after all—Extrinsic rewards, task interest, and self-determination: Reply to Eisenberger, Pierce, and Cameron (1999) and Lepper, Henderlong, and Gingras (1999). *Psychol. Bull.* **1999**, *125*, 692–700. [[CrossRef](#)]
11. Kazuyoshi, T.; Madoka, M.; Yousuke, O.; Keiko, M.; Hiroki, M.; Kou, M.; Keigo, S.; Takashi, H.; Kenji, M.; Kazuyuki, N. Impaired prefrontal activity to regulate the intrinsic motivation-action link in schizophrenia. *NeuroImage Clin.* **2017**, *16*, 32–42. [[CrossRef](#)]

12. Ma, Q.; Jin, J.; Meng, L.; Shen, Q. The dark side of monetary incentive: How does extrinsic reward crowd out intrinsic motivation. *NeuroReport* **2014**, *25*, 194–198. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Jin, J.; Yu, L.; Ma, Q. Neural Basis of Intrinsic Motivation: Evidence from Event-Related Potentials. *Comput. Intell. Neurosci.* **2015**, *2015*, 1–6. [\[CrossRef\]](#) [\[PubMed\]](#)
14. de Vicente, A.; Pain, H. Motivation diagnosis in intelligent tutoring systems. In Proceedings of the Fourth International Conference on ITS, San Antonio, TX, USA, 16–19 August 1998; Barry, P., Goettl, H.M., Half, C.L.R., Valerie, J.S., Eds.; Springer: Berlin/Heidelberg, Germany, 1998; pp. 86–95.
15. De Vicente, A.; Pain, H. Informing the detection of the students' motivational state: An empirical study. In *International Conference on Intelligent Tutoring Systems*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 933–943.
16. Organero, M.M.; Merino, P.J.M.; Delgado-Kloos, C. Student Behavior and Interaction Patterns With an LMS as Motivation Predictors in E-Learning Settings. *IEEE Trans. Educ.* **2010**, *53*, 463–470. [\[CrossRef\]](#)
17. Chattopadhyay, S.; Zary, L.; Quek, C.; Prasad, D.K. Motivation detection using EEG signal analysis by residual-in-residual convolutional neural network. *Expert Syst. Appl.* **2021**, *184*, 115548. [\[CrossRef\]](#)
18. Feldman Barrett, L.; Russell, J.A. Independence and bipolarity in the structure of current affect. *J. Personal. Soc. Psychol.* **1998**, *74*, 967–984. [\[CrossRef\]](#)
19. Nasoz, F.; Alvarez, K.; Christine, L.; Finkelstein, N. Emotion Recognition from Physiological Signals for User Modeling of Affect. In Proceedings of the 9th International Conference on User Mode, Pittsburg, PA, USA, 22–26 June 2003.
20. Nardelli, M.; Valenza, G.; Greco, A.; Lanata, A.; Scilingo, E.P. Recognizing Emotions Induced by Affective Sounds through Heart Rate Variability. *IEEE Trans. Affect. Comput.* **2015**, *6*, 385–394. [\[CrossRef\]](#)
21. Bradley, M.M.; Lang, P.J. Measuring emotion: Behavior, feeling, and physiology. In *Cognitive Neuroscience of Emotion*; Lane, R.D., Nadel, L., Eds.; Oxford University Press: Oxford, UK, 2000; pp. 242–276.
22. Kurita, K.; Yonezawa, T.; Kuroshima, M.; Tsumura, N. Non-Contact Video Based Estimation for Heart Rate Variability Spectrogram using Ambient Light by Extracting Hemoglobin Information. In Proceedings of the Color and Imaging Conference, Darmstadt, Germany, 19–23 October 2015. [\[CrossRef\]](#)
23. Selcan, K.B.; Alper, K.U.; Efnan, S.G.; Semih, E.; Serkan, G.; Bilginer, G.M. A survey on ECG analysis. *Biomed. Signal Process. Control.* **2018**, *43*, 216–235. [\[CrossRef\]](#)
24. Malac, A.A.; Saiful, I.; Saad, A.A.; Ahmed, S.B. Diagnostic Features and Potential Applications of PPG Signal in Healthcare: A Systematic Review. *Healthcare* **2022**, *10*, 547. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Yu, S.; Nitish, T. Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging. *IEEE Trans. Bio-Med. Eng.* **2015**, *63*, 463–477. [\[CrossRef\]](#)
26. Shahid, H.; Butt, A.; Aziz, S.; Khan, M.U.; Naqvi, S.Z.H. Emotion Recognition System featuring a fusion of Electrocardiogram and Photoplethysmogram Features. In Proceedings of the 14th International Conference on Open Source Systems and Technologies, Lahore, Pakistan, 16–17 December 2020. [\[CrossRef\]](#)
27. Lee, M.S.; Lee, Y.K.; Pae, D.S.; Lim, M.T.; Kim, D.W.; Kang, T.K. Fast Emotion Recognition Based on Single Pulse PPG Signal with Convolutional Neural Network. *Appl. Sci.* **2019**, *9*, 3355. [\[CrossRef\]](#)
28. Poh, M.; McDuff, D.J.; Picard, R.W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt. Express* **2010**, *18*, 10762–10774. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Okubo, T.; Ono, K.; Tsumura, N. Improving driving ability using biofeedback by monitoring the mental situation by RGB camera. In Proceedings of the SPIE BiOS, San Francisco, CA, USA, 16 March 2023. [\[CrossRef\]](#)
30. Tsumura, N.; Ojima, N.; Sato, K.; Shiraishi, M. Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Trans. Graph.* **2003**, *22*, 770–779. [\[CrossRef\]](#)
31. Tanaka, S.; Tsumura, N. Improved analysis for skin color separation based on independent component analysis. *Artif. Life Robot.* **2019**, *25*, 159–166. [\[CrossRef\]](#)
32. Atkinson, J.W. Motivational determinants of risk-taking behavior. *Psychol. Rev.* **1957**, *64*, 359–372. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Kumar, M.; Veeraraghavan, A.; Sabharwal, A. Distance-PPG: Robust Non-Contact Vital Signs Monitoring Using a Camera. *Biomed. Opt. Express* **2015**, *6*, 1565–1588. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Tarvainen, M.P.; Ranta-aho, P.O.; Karjalainen, P.A. An advanced detrending method with application to HRV analysis. *IEEE Trans. Biomed. Eng.* **2002**, *49*, 172–175. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Shaffer, F.; Ginsberg, J.P. An Overview of Heart Rate Variability Metrics and Norms. *Front. Public Health* **2017**, *5*, 258. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Ekman, P.; Friesen, W.V. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*; Consulting Psychologists Press: Washington, DC, USA, 1978.
37. Cohn, J.F.; Ambadar, Z.; Ekman, P. *Handbook of Emotion Elicitation and Assessment*; Oxford University Press: Oxford, UK, 2007; pp. 203–221.
38. Baltrusaitis, T.; Robinson, P.; Morency, L.P. OpenFace: An open source facial behavior analysis toolkit. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, Lake Placid, NY, USA, 7–10 March 2016. [\[CrossRef\]](#)
39. Ekman, P.; Friesen, W.V. Constants across cultures in the face and emotion. *J. Personal. Soc. Psychol.* **1971**, *17*, 124–129. [\[CrossRef\]](#) [\[PubMed\]](#)

40. Ekman, P.; Friesen, W.V.; Hager, J.C. Facial Action Coding System. In *Manual and Investigator's Guide*; Consulting Psychologists Press: Salt Lake City, UT, USA, 2002.
41. scikit-learn 1.3.0 documentation. Sequential Feature Selector. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SequentialFeatureSelector.html (accessed on 27 June 2024).
42. scikit-learn 1.3.0 documentation. Recursive Feature Elimination. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html (accessed on 27 June 2024).
43. sklearn.linear_model.LogisticRegression. Available online: https://scikit-learn.org/1.3/modules/generated/sklearn.linear_model.LogisticRegression.html (accessed on 27 June 2024).
44. sklearn.ensemble.RandomForestClassifier. Available online: <https://scikit-learn.org/1.3/modules/generated/sklearn.ensemble.RandomForestClassifier.html> (accessed on 27 June 2024).
45. sklearn.svm.SVC. Available online: <https://scikit-learn.org/1.3/modules/generated/sklearn.svm.SVC.html> (accessed on 27 June 2024).
46. Appel, M.L.; Berger, R.D.; Saul, J.P.; Smith, J.M.; Cohen, R.J. Beat to beat variability in cardiovascular variables: Noise or music? *J. Am. Coll. Cardiol.* **1989**, *14*, 1139–1148. [[CrossRef](#)]
47. Malik, M.; Bigger, J.T.; Camm, A.J.; Kleiger, R.E.; Malliani, A.; Moss, A.J.; Schwartz, P.J. Heart rate variability: Standards of measurement, Physiological Interpretation, and Clinical Use. *Eur. Heart J.* **1996**, *17*, 354–381. [[CrossRef](#)]
48. Immanuel, S.; Teferri, M.N.; Baumert, M.; Bidargaddi, N. Heart Rate Variability for Evaluating Psychological Stress Changes in Healthy Adults: A Scoping Review. *Neuropsychobiology* **2023**, *82*, 187–202. [[CrossRef](#)] [[PubMed](#)]
49. Schildkraut, J.J. The catecholamine hypothesis of affective disorders: A review of supporting evidence. *Am. J. Psychiatry* **1965**, *122*, 509–522. [[CrossRef](#)] [[PubMed](#)]
50. Di Domenico, S.I.; Ryan, R.M. The Emerging Neuroscience of Intrinsic Motivation: A New Frontier in Self-Determination Research. *Front. Hum. Neurosci.* **2017**, *11*, 1–14. [[CrossRef](#)] [[PubMed](#)]
51. Vandercammen, L.; Hofmans, J.; Theuns, P. Relating Specific Emotions to Intrinsic Motivation: On the Moderating Role of Positive and Negative Emotion Differentiation. *PLoS ONE* **2014**, *9*, 1–22. [[CrossRef](#)] [[PubMed](#)]
52. Csikszentmihalyi, M.; LeFevre, J. Optimal experience in work and leisure. *J. Pers. Soc. Psychol.* **1989**, *5*, 815–822. [[CrossRef](#)] [[PubMed](#)]
53. Nakamura, J.; Tse, D.C.K.; Shankland, S. Flow: The experience of intrinsic motivation. In *The Oxford Handbook of Human Motivation*, 2nd ed.; Ryan, R.M., Ed.; Oxford University Press: Oxford, UK, 2019; pp. 169–185.
54. Cheng, Y.; Mukhopadhyay, A.; Williams, P. Smiling Signals Intrinsic Motivation. *J. Consum. Res.* **2020**, *46*, 915–935. [[CrossRef](#)]
55. Ewa, M.E.; Daniel, M.; Ashok, V. A Meta-Analysis of the Impact of Skin Type and Gender on Non-contact Photoplethysmography Measurements. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020. [[CrossRef](#)]
56. Goodfellow, I.J.; Shlens, J.; Szegedy, C. Explaining and harnessing adversarial examples. *arXiv* **2014**, arXiv:1412.6572.
57. Baia, A.E.; Biondi, G.; Franzoni, V.; Milani, A.; Poggioni, V. Lie to Me: Shield Your Emotions from Prying Software. *Sensors* **2022**, *22*, 967. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.