Check for updates

OPEN ACCESS

EDITED BY Elias Ebrahimzadeh, University of Tehran, Iran

REVIEWED BY Siamak Aram, Harrisburg University of Science and Technology, United States Joanne Harris, University of South Australia Online, Australia

*CORRESPONDENCE Marco Bilucaglia Imarco.bilucaglia@studenti.iulm.it

RECEIVED 30 November 2023 ACCEPTED 28 February 2024 PUBLISHED 20 March 2024

CITATION

Bilucaglia M, Zito M, Fici A, Casiraghi C, Rivetti F, Bellati M and Russo V (2024) I DARE: IULM Dataset of Affective Responses. *Front. Hum. Neurosci.* 18:1347327. doi: 10.3389/fnhum.2024.1347327

COPYRIGHT

© 2024 Bilucaglia, Zito, Fici, Casiraghi, Rivetti, Bellati and Russo. This is an open-access article distributed under the terms of the **Creative Commons Attribution License (CC BY)**. The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

I DARE: IULM Dataset of Affective Responses

Marco Bilucaglia^{1,2}*, Margherita Zito^{1,2}, Alessandro Fici^{1,2}, Chiara Casiraghi^{1,2}, Fiamma Rivetti¹, Mara Bellati¹ and Vincenzo Russo^{1,2}

¹Behaviour and Brain Lab, Neuromarketing Research Center, Università IULM, Milan, Italy, ²Department of Business, Law, Economics and Consumer Behaviour "Carlo A. Ricciardi", Università IULM, Milan, Italy

KEYWORDS

Affect Detection, affective database, Neuromarketing, Consumer Neuroscience, machine learning, EEG, physiological signals

1 Introduction

1.1 Consumer neuroscience and neuromarketing

Consumer Neuroscience and Neuromarketing apply neuroscience tools and methodologies to investigate consumer behavior (Karmarkar and Plassmann, 2017; Lim, 2018), addressing the limitations of traditional marketing techniques (Hakim and Levy, 2018). Although often used interchangeably, these two terms have different meanings: Consumer Neuroscience, which is more theoretical/academic, is considered as the scientific foundation of Neuromarketing, a more practical and commercially oriented discipline (Nemorin, 2016; Ramsøy, 2019).

Established in the early 2000s (Smith, 2002; Yoon et al., 2006), Neuromarketing and Consumer Neuroscience have become increasingly popular topics among researchers. According to a recent review (Casado-Aranda et al., 2023), as of 2022, 862 articles were published in international journals belonging to psychology (28.45%), business (24.75%), communication (24.08%), and neuroscience (15%), along with management, information technology, and behavioral science (7.72%). The scientific production increased annually by 7.5% and involved 2,354 contributors (3.83 average co-authors per study) with 24.97% of international collaborations.

Consumer Neuroscience and Neuromarketing collect neurophysiological and biometric signals (decoded into "neurometrics") from subjects during the exposure to marketing stimuli/tasks, in order to assess the underlying cognitive and emotional processes (Cherubino et al., 2019).

The most frequently recorded signals include electroencephalogram—EEG (56.14%), eye tracking—ET (10.53%), skin conductance—SC, cardiac activity (8.77%), and electromyogram - EMG (1.75%) (Rawnaque et al., 2020). Facial expressions are also included in the standard set of Neuromarketing and Consumer Neuroscience measures (Mancini et al., 2023). Typical stimuli include video advertisements (Zito et al., 2021; Russo et al., 2022, 2023), radio commercials (Russo et al., 2020), product packaging (Russo et al., 2021), static creativities (Ciceri et al., 2020), and websites (Bengoechea et al., 2023). Common tasks include the buying processes (Cherubino et al., 2017), in-store experiences (Balconi et al., 2021), and food choices (Stasi et al., 2018).

Decision-making, memory, attention, and approach-withdrawal motivation are the main assessed psychological processes (Cherubino et al., 2019), but emotions, above all, have attracted the greatest scientific interest (Casado-Aranda et al., 2023), probably due to their well-established role in mediating several cognitive functions (Pessoa, 2008; Lerner et al., 2015; Soodan and Pandey, 2016; Tyng et al., 2017).

Neurometrics were originally intended as invariant and highly selective "biomarkers", as they derived from a large number of experimental neuroimaging studies (Plassmann et al., 2012). However, they have soon been called into question due to their correlational (rather than causal) nature (Plassmann et al., 2015; Ramsøy, 2019), which leads to the reverse inference fallacy (Lee et al., 2017), a common albeit serious problem in neuroscience (Poldrack, 2006). Despite Machine Learning models have been suggested as a reliable solution for the reverse inference issue, they still remain largely ignored (Nathan and Pinal, 2017). Indeed, as of 2019, only 35.1% of Neuromarketing and Consumer Neuroscience studies employed these techniques (Rawnaque et al., 2020).

1.2 Affect Detection

Affect Detection (AD) aims to design automatic systems to recognize human emotions from neurophysiological, biometric, and behavioral signals (Calvo and D'Mello, 2010). Initially developed within the Affective Computing, a Computer Science subbranch established in the late 1990s (Picard, 1995, 1997), AD had a far-reaching impact, influencing several research contexts, such as Transportation, Defense, and Secutity, Human Resources, Education, Mental Health and even Market Research (Hernandez et al., 2021).

Since its original formulation, AD has been treated as a Machine Learning problem (Picard, 1995) based on a model f trained to predict the affective state s from a set of neurophisiological and biometric variables $\mathbf{m} = (m_1, ..., m_m)^T \in \mathbb{R}^m$, categorized into "modalities" (Picard, 1997). Depending on the chosen emotional theory (Gunes and Pantic, 2010), the emotional state s is modeled as a discrete or continuous quantity, which, in turn, determines whether the problem will be treated as a regression or a classification, respectively. A comprehensive discussion on the Machine Learning theory applied to the emotion recognition field goes beyond the scope of this article. Interested readers can refer to introductory tutorials, such as Zhang et al. (2020) and Can et al. (2023).

The classification approach is preferred, even in the case of a continuous emotional model, where the dimensions are early categorized (e.g., low, medium and high). Despite 63% of the studies referred to the continuous emotional models, 92% of them adopted classifiers (D'Mello and Kory, 2015), regardless of the modalities (90% with the voice and 92% with physiological signals) (Imani and Montazer, 2019) and within the wearable context (98%) (Schmidt et al., 2019). Furthermore, a more recent survey that examined more than 380 AD papers published between 2000 and 2020 did not report any regression methods (Wang et al., 2022).

Common modalities include facial expressions, voice, gaze movements, and physiological signals related to either the Peripheral (PNS) or the Central Nervous System (CNS) (Calvo and D'Mello, 2010). Signals from the PNS include EMG, SC, Electrocardiogram (EEG), Photoplestimogram (PPG), and Electrooculogram (EOG). Signals from CNS include EEG, functional Magnetic Resonance Imaging (fMRI), and functional Near-Infrared Spectroscopy (fNIRS) (Shu et al., 2018; Larradet et al., 2020). Voice and facial expressions are the most popular, TABLE 1 Qualitative comparison between I DARE and the eight most popular physiological-based datasets in terms of modalities and stimuli type.

Dataset	Modalities	Stimuli
AMIGOS	EEG, ECG, EDA	Video clips
ASCERTRAIN	EEG, ECG, EDA	Movie clips
BIO-VID-EMO DB	ECG, EMG, SC	Film clips
DEAP	EEG, EDA, EMG, PPG, EOG, Resp	Music videos
DREAMER	EEG, ECG	Film clips
MAHNOB-HCI	EEG, ECG, SC, Resp, Temp, ET	Video clips
MPED	ECG, EEG, PPG, SC, Resp	Video clips
SEED	EEG	Film clips
I DARE	EEG, EMG, SC, PPG, ET	Images

varying from 82% and 77% (D'Mello and Kory, 2015) to 21% and 54% (Wang et al., 2022), respectively. Physiological signals are adopted in ~16% of the models overall (Wang et al., 2022): at least 10% from the PNS and 5% from the CNS (D'Mello and Kory, 2015). Wearable applications draw a completely different scenario: the modalities include 100% PNS signals, 15% CNS, only 2% facial expressions, and never the voice (Schmidt et al., 2019). Multimodal models (i.e., based on at least two modalities) vary from 85% (D'Mello and Kory, 2015) to 93% (Schmidt et al., 2019).

Affective datasets are collections of modalities annotated in terms of the corresponding affective states and are used to train AD models (D'Mello et al., 2017). The modalities can be either extracted from existing data (e.g., facial expressions from video clips) or newly recorded during experiments (e.g., EEG data from a sample of subjects). Furthermore, affective states can be induced by external stimuli or emerge naturally, as well as posed or spontaneously emitted (D'Mello and Kory, 2015; Zhao et al., 2021).

In the literature, multiple affective datasets that vary largely in terms of modalities, labels, and origin of the affective state are found. We have taken into account a representative sample of 58 datasets, gathered from two reviews (Siddiqui et al., 2022; Ahmed et al., 2023) and three recent research studies (Chen et al., 2022; Saganowski et al., 2022; Pant et al., 2023). Accordingly, the number of modalities varies from 1 to 7 (M = 2.28, SD = 1.24) and include 71.2% the face, 54.2% the voice, 23.7% EEG, 20.3% ECG, 17% PPG, 6.8% EMG, and 6.8% ET. Overall, 81% of the datasets contain newly recorded modalities. The affective states, that are mostly labeled according to a discrete model (70.7%), are mainly induced (69%) and spontaneous (63%).

A recent survey (Ahmad and Khan, 2022) compared the stimuli, modalities, and sample size of the eight most popular datasets for the training of physiological-based AD models: AMIGOS (Miranda-Correa et al., 2018), ASCERTRAIN (Subramanian et al., 2016), BIO-VID-EMO DB (Zhang et al., 2016), DEAP (Koelstra et al., 2011), DREAMER (Katsigiannis and Ramzan, 2017), MAHNOB-HCI (Soleymani et al., 2011), MPED (Song et al., 2019), and SEED (Zheng et al., 2017). The stimuli, varying from 10 to 40 (M = 23.5, SD = 9.943) are all induced and spontaneous. The number of modalities, varying from 1 to 6 (M = 3.38, SD = 1.60), includes EEG (87.5%), ECG (75%), SC (37.5%), respiration (37.5%), PPG (25%), EMG (25%), peripheric temperature—Temp (12.5%), and ET (12.5%). The sample size ranges from 15 to 86 (M = 38, SD = 23.43). The "dataset dimension", defined as the product between stimuli, modalities, and sample sizes, ranges from 225 to 7,680 (M = 3,257.88, SD = 2,581.12). Finally, the stimuli are all dynamic, including video clips (37.5%), film clips (37.5%), movie clip (12.5%), and music video (12.5%). The features of the eight most popular datasets (plus the proposed one) are presented in Tables 1, 2.

For each dataset, we conducted statistical analyses using JASP (Love et al., 2019), to explore relationships between the number of modalities, stimuli, and sample size. Additionally, we performed sensitivity analyses using G*Power (Faul et al., 2007) to estimate, from the sample size and the number of stimuli, the minimum detectable effect size. No significant correlations between the number of modalities and subjects (Spearman's correlation: $\rho =$ 0.270, p = 0.518), modalities and stimuli (Spearman's correlation: $\rho = 0.700, p = 0.053$), and subjects and stimuli (Spearman's correlation: $\rho = 0.084, p = 0.843$) emerged, suggesting that the dataset-specific characteristics are not bounded by technological or experimental constrains. The sensitivity analysis, which was based on the repeated measures ANOVA model (stimulus ID as within factor, $\alpha = 0.05, 1 - \beta = 0.95, r = 0.5, \epsilon = 1$), showed an average minimum detectable effect size of $f = 0.157 \pm 0.051$ ranging from 0.095 to 0.253, interpretable as "small" to "medium" (Cohen, 2013).

1.3 I DARE dataset

We contend that the diffusion of AD models within Neuromarketing and Consumer Neuroscience could be increased by addressing the limitations of existing datasets. We believe that the strictest one is the use of videos as a unique elicitation method, which may lead to AD models that are unsuitable for the evaluation of static or short-term stimuli (e.g., static creativities, product packaging, or landing pages), as well as for a dynamic classification (Bilucaglia et al., 2021) of multi-frame stimuli (e.g., video commercials). A less severe (but still present) limitation is the varying number of stimuli that affects the minimum detectable effect size and, in turn, the reliability and generalizability of the results (Funder and Ozer, 2019). Finally, despite their popularity in Consumer Neuroscience and Neuromarketing, modalities such as ET, SC, PPG and EMG are not as widely represented.

We present I DARE, the <u>IULM Dataset of Affective REsponses</u>, a multimodal (five modalities: EEG, SC, PPG, EMG, and ET) physiological affective dataset gathered from 63 participants, based on 32 images as elicting stimuli, corresponding to a dimension of $5 \times 32 \times 63 = 10,080$. A sensitivity analysis, performed using G*Power under the same previously described conditions, showed a minimum detectable effect size of f = 0.095, considered as "small". As shown in Table 2, I DARE exceeds existing dataset in three key aspects: it is based on static pictures, it has the highest dimension, and it allows the lowest minimum detectable effect size.

2 Materials and methods

2.1 Stimuli selection

We extracted the eliciting stimuli from both the IAPS (Lang et al., 2008) and OASIS (Kurdi et al., 2017), two standard datasets of realistic pictures that have been annotated in the valence and arousal dimensions by ~100 independent raters. The annotations are provided as mean (*m*) and standard deviation (*s*) across the raters, for both the valence (m_v , s_v) and the arousal (m_a , s_a) dimensions.

For each dataset, we performed the stimuli selection through a seven-step semi-automatic procedure. Steps 1–5 (automatic) were based on the geometrical properties of the stimuli in the two-dimensional Valence-Arousal space, which were considered either as points $\mathbf{x} = (m_v, m_a)^T$ or gaussian vectors $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\mu} = (m_v, m_a)^T$ and $\boldsymbol{\Sigma} = \begin{pmatrix} s_v & 0 \\ 0 & s_a \end{pmatrix}$. Steps 6, 7 were based on the ratings provided by nine independent judges.

The selection process yielded 32 stimuli (16 per dataset), belonging to each of the 4 quadrants of the Valence-Arousal space (HVLA - High Valence High Arousal, HVLA - High Valence Low Arousal, LVHA—Low Valence High Arousal, and LVLA - Low Valence Low Arousal). The whole procedure is shown in the https://figshare.com/projects/I_DARE/186558 (Stimulus_Selection.pdf file).

2.2 Sample population and experimental protocol

Sixty-three people (24 men) with age ranging from 20 to 30 years (M = 22.89, SD = 2.38) were enrolled in the study. Statistical analyses performed using JASP showed that the sample was not significantly gender unbalanced in terms of proportions (Binomial test: Male = 38.1%, p = 0.077) and mean age (Mann–Whitney U test: W = 3995, p = 0.325). The sample characteristics are shown in https://figshare.com/projects/I_DARE/186558 (Sample.csv file).

The protocol was approved by the Ethics Committee of the Università IULM, and an informed consensus was obtained from each participant.

Each subject sat on a chair placed in front of a 23.8'' (527 × 296.5mm²) FlexScan EV2451 (Eizo, KK) monitor located in a 7 × 3m² room. At the beginning of the experiment, the subject performed a 60 s-long eye-closed baseline and a 120 slong neutral baseline (white fixation dot on a black background). This was preceded by the baseline instructions and followed by the experiment description. Then, 32 emotional blocks were presented in a random order. Each block consisted of a black screen (5 s), followed by the emotional stimulus (5 s) and the valence/arousal evaluation, performed by means of a nine-point SAM (Bradley and Lang, 1994). The description and duration of the stimuli are shown in https://figshare.com/projects/I_DARE/186558 (Stimuli_Specification.csv file).

The quality of the signals was visually checked before starting the recording, and the contact impedance of the EEG electrodes was reduced below $10k\Omega$ (Sinha et al., 2016) by means of a scrub cream (Neurprep by Spes Medica, S.p.A.) and a conductive gel (Neurgel by Spes Medica, S.p.A.).

Dataset	# Modalities	# Stimuli	# Subjects	Dimension	Cohen's f
AMIGOS	3	16	40	1,920	0.149
ASCERTRAIN	3	36	58	6,264	0.095
BIO-VID-EMO DB	3	15	86	3,870	0.104
DEAP	6	40	32	7,680	0.125
DREAMER	2	18	23	828	0.191
MAHNOB-HCI	5	20	27	2,700	0.170
MPED	4	28	23	2,576	0.166
SEED	1	15	15	225	0.253
I DARE	5	32	63	10,080	0.095

TABLE 2 Quantitative comparison between I DARE and the eight most popular physiological-based datasets in terms of number of modalities, stimuli, and subjects, as well as dimension and minimum detectable effect size (Cohen's f).

2.3 Instrumentation

We recorded the EEG signal using the NVX-52 device (Medical Computer Systems, Ltd.) at a sample frequency of $f_s = 1kHz$. Forty passive wet Ag / AgCl electrodes were placed at standard locations (Russo et al., 2023) of the 10-10 system (Nuwer, 2018) and referenced to the left mastoid (M1) by means of 1 Ag/AgCl passive adhesive patch. We recorded the SC and PPG signals using the GSRSens and FpSens (Medical Computer Systems, Ltd.) sensors, respectively, both connected to the auxiliary inputs of the NVX-52. We placed two passive dry Ag/AgCl electrodes of the GSRSens on the index and middle finger from the non-dominant hand and the FpSens on the ring finger from the same hand. We recorded the EMG signal using two passive Ag/AgCl adhesive patches placed over the left zygomaticus major and corrugators supercilii muscles. The electrodes, connected to the bipolar inputs of the NVX-52, were referenced to M1. The recordings were controlled by NeoRec software (Medical Computer Systems, Ltd.). We recorded the ET signal using the Tobii Spectrum (Tobii, LLC) device at a sample frequency of $f_s = 150Hz$. The collected measures included the horizontal and vertical coordinates of the right and left eyes $(x_R, y_R, x_R, \text{ and } y_L, \text{ respectively})$, expressed in *mm* with respect to the upper-left corner of the screen.

We used the iMotions software (iMotions, A/V) to deliver the stimuli and to collect the eye tracking data and the SAM responses. At the beginning of the experiment, iMotions generated a TTL pulse that was fed into the digital inputs of the NVX-52 using the ESB - EEG Synchronization Box (Bilucaglia et al., 2020). This served to perform an off-line synchronization between the recorded data and the stimuli timestamps.

2.4 Data processing

We aligned the recorded data with the TTL pulse and appended the stimuli, according to the iMoions-generated timestamps. Then, we cleaned the data from noise and artifacts, as described below. The operations were performed within the Matlab (The Mathworks, Inc.) environment, and the processed data were saved as .mat files. A comprehensive discussion on the biomedical data processing goes beyond the scope of this article. Interested readers can refer to introductory textbooks, such as Blinowska and Żygierewicz (2021) and Rangayyan and Krishnan (2024).

2.4.1 EEG

We processed the EEG data using the EEGLab toolbox (Delorme and Makeig, 2004). We applied a re-reference to the linked earlobes, a resample to $f_S = 512Hz$, a band-pass filter (0.1 – 40Hz, 4th order zero-phased Butterworth filter), and an adaptive filter (cleanline plug-in), to remove the sinusoidal noise (50Hz, 100Hz). We applied the artifact Subspace Reconstruction (cutoff parameter k = 20) to correct large non-stationary artifacts (Chang et al., 2018). Then, we performed the Independent Component Analysis (ICA) decomposition trough the RunICA algorithm (Akhtar et al., 2012) on a copy of the original data that have been band-pass filtered (1 - 30Hz, 4th order zerophased Butterworth filter) and down-sampled ($f_s = 100Hz$) (Luck, 2022). The obtained ICA weights were then saved and multiplied to the original data, to get the ICs. We identified artifactual ICs through the neural-net classifier ICLabel (Pion-Tonachini et al., 2019) as those with probability of "non-brain" P{!brain}>0.9, and we set them to zero and back-projected them to the original space. Finally, we re-referenced the artifactfree data to the approximatively-zero potential REST (Dong et al., 2017).

2.4.2 SC

We applied a resample at $f_S = 32Hz$ and a band-pass filter (0.001 – 0.35Hz, 4th order zero-phased Butterworth filter). We identified artifactual points as those exceeding two fixed thresholds (amplitude: 0.05–60 μ S, rate of change: $\pm 8 \mu$ S/s) and deleted all the neighbour within a centered 5*s*—long window (Kleckner et al., 2018). Finally, we reconstructed the deleted points by means of a linear interpolation applied on the whole data.

2.4.3 PPG

We applied a resample at $f_S = 32Hz$, a band-pass filter (0.001 – 0.35Hz, 2*nd* order zero-phased Butterworth filter), and a filter based on the Hilbert's transform, to attenuate slow drifts and motion artifacts (Dall'Olio et al., 2020).

2.4.4 EMG

We applied a notch filter (50*Hz*, 100*Hz*, q = 40, zero-phased IIR filter), to remove the power-line noise and a band-pass filter (10-400*Hz*, 4*th* order zero-phased Butterworth filter), to attenuate low-frequency motion artifacts (Merletti and Cerone, 2020).

2.4.5 ET

We corrected temporal jitters by re-aligning the timestamps with the equipment-designed sample frequency (fs = 150Hz) and linearly interpolating the missing values (Liu et al., 2018). Then, we filtered the coordinates x_R , y_R , x_L , and y_L using the MFLWMD algorithm (window length 140*ms*, threshold g = 280ms), to reduce outliers and missing data (Gucciardi et al., 2022). Finally, we computed the horizontal and vertical gaze positions as, respectively, $g_X = (x_R + x_L)/2$ and $g_Y = (y_R + y_L)/2$ (Kar, 2020).

3 Dataset description

I DARE includes the output of the stimuli selection process, the SAM evaluations, and the processed modalities.

3.1 Stimuli selection

The selected stimuli obtained a percentage agreement ranging from 56.56% to 100% (M = 79.20%, SD = 10.50%). Statistical analyses performed using JASP did not show any significant main effect for the factors "Database" $[F_{(1, 24)} = 0.143, p = 0.709]$ and "Quadrant" $[F_{(3, 24)} = 2.381, p = 0.095]$, as well as for the "Database×Quadrant" $[F_{(3, 24)} = 0.048, p = 0.986]$ interaction. Additionally, the agreement did not significantly correlate with the mean (m) and standard deviations (s) of the overall ratings in both the valence $(\rho_m = 0.194, p = 0.288; \rho_s = -0.175, p = 0.339)$ and the arousal $(\rho_m = -0.142, p = 0.439; \rho_s = 0.109, p = 0.553)$ dimensions. The results of the stimuli selection process are shown in https://figshare.com/projects/I_DARE/186558 (Agreement_Raters.csv file).

3.2 SAM evaluations

We coded the subjects' evaluations (SAM numerical scores) into labels corresponding to the four quadrants (HVLA, LVLA, HVHA, and LVHA), considering the mid-point value 5 as threshold for the low/high discrimination. We measured the overall agreement within the subjects (*WS*) and between the subjects and the gold standard (*GS*) given by the selection process by means of the Fleiss' kappa (Fleiss, 1971). The analyses, performed using JASP, showed values of $\kappa_{WS} = 0.406$ and $\kappa_{GS} = 0.409$,

both intepreted as a "fair agreement" (Landis and Koch, 1977). The SAM evaluations are shown in https://figshare.com/projects/I_DARE/186558 (Valence_SAM.csv, Arousal_SAM.csv and Quadrants_SAM.csv files).

3.3 Modalities

We grouped the processed data based on their sample frequency and stored them into separated folders (EEG, EMG, SC&PPG, and ET) containing 63 .mat files each (1 per subject). The files consist of Matlab's "structure arrays" with the following fields:

- subject: subject's unique identifier
- Fs: sample frequency
- channels: channels' name
- channels_type: modality (EEG, SC, PPG, and ET)
- channels_unit: channels units of measurement (μV μS, a.u., mm)
- data: data matrix (rows: channels, columns: sample points)
- time: vector of sample points
- event_id: stimuli labels
- event_begin: stimuli onsets as positions within the time vector
- even_end: stimuli offsets as positions within the time vector

4 Conclusion

Emotions are the most commonly explored aspect within Consumer Neuroscience and Neuromarketing. Their assessment has been mainly performed by mean of neurometrics, an approach often questioned due to the potential reverse inference problem. By contrast, Machine Learning methods (i.e., AD models), that have been proposed as a reliable solution to the issue, are still under-represented.

Existing physiological datasets rely on dynamic eliciting stimuli, making them unsuitable to train AD models for the evaluation of images or the dynamic frame-by-frame classification of videos. Furthermore, the available data sets are characterized by a wide-varying number of subjects, stimuli, and modalities, negatively affecting their reliability and generalizability.

To fill these gaps, we created I DARE (the <u>IULM</u> <u>D</u>ataset of <u>Affective <u>RE</u>sponses), a multimodal physiological dataset that outperforms existing ones in terms of dimension (5 modalities \times 32 stimuli \times 63 subjects) and minimum detectable effect size (f = 0.095, small). We selected the stimuli (static pictures) from the two standard databases (IAPS and OASIS) through a semiautomatic procedure, to reduce the selection biases. We obtained "fair" agreements within subjects and between subjects and the original scores ($\kappa_{WS} = 0.406$ and $\kappa_{GS} = 0.409$, respectively).</u>

We must acknowledge some limitations of I DARE. First, while the sample was not significantly gender unbalanced, it mainly consisted of young subjects (i.e., \leq 30 years old). This could potentially impact the sensitivity of trained AD models toward age differences. Including new data from older subjects

would not only help to address the imbalance in the dataset but also increase its size and, consequently, its representativity. Second, the experimental paradigm, which was chosen to minimize confounding effects and biases, may also have affected the ecological validity of the collected data. As a result, the applicability of trained AD models in real-life situations may be limited. To examine this hypothesis, future experiments outside the laboratory, using natural stimuli and wearable devices, should be carried out.

Nevertheless, we believe that I DARE could be helpful in several applications. It could be used for the emotional assessment of static creativities (e.g., product packaging, landing pages) and the dynamic classifications of video commercials in Consumer Neuroscience and Neuromarketing researches, as well as to train AD models for Affective Computing and Neuroergonomics applications. Additionally, it could serve as a benchmark for the assessment of the emotional selectivity and invariance of existing neurometrics, as well as to study the physiology of emotions in the Psychology and Cognitive Neuroscience fields.

To foster its adoption by a larger audience of scientists, I DARE has been made publicly available at https://figshare.com/projects/I_DARE/186558.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://figshare.com/projects/I_DARE/186558.

Ethics statement

The studies involving humans were approved by Ethics Committee of Università IULM. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

References

Ahmad, Z., and Khan, N. (2022). A survey on physiological signal-based emotion recognition. *Bioengineering* 9:688. doi: 10.3390/bioengineering9110688

Ahmed, N., Aghbari, Z. A., and Girija, S. (2023). A systematic survey on multimodal emotion recognition using learning algorithms. *Intell. Syst. Appl.* 17:200171. doi: 10.1016/j.iswa.2022.200171

Akhtar, M. T., Jung, T.-P., Makeig, S., and Cauwenberghs, G. (2012). "Recursive independent component analysis for online blind source separation," in 2012 IEEE International Symposium on Circuits and Systems (Seoul: IEEE), 2013–2016.

Balconi, M., Venturella, I., Sebastiani, R., and Angioletti, L. (2021). Touching to feel: brain activity during in-store consumer experience. *Front. Psychol.* 12:653011. doi: 10.3389/fpsyg.2021.653011

Bengoechea, A. G., Ruiz, F. J. G., Herrera, M. H., and Crespo, R. A. (2023). Neuromarketing and e-commerce: analysis of over the top platform homepages. *Int. J. Serv. Operat. Inf.* 12, 253–266. doi: 10.1504/IJSOI.2023.132354

Author contributions

MBi: Conceptualization, Data curation, Formal analysis, Methodology, Writing—original draft, Writing—review & editing. MZ: Supervision, Writing—review & editing. AF: Methodology, Writing—original draft. CC: Writing—original draft. FR: Investigation, Methodology, Writing—review & editing. MBe: Investigation, Methodology, Writing—review & editing. VR: Supervision, Writing—review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would express their gratitude to their former colleague, Riccardo Circi, Ph.D., for his valuable support during the initial stages of experiment design and data acquisition. The authors acknowledge Chiara Falceri, B.Sc., for the assistance in language editing and proofreading.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Bilucaglia, M., Duma, G. M., Mento, G., Semenzato, L., and Tressoldi, P. (2021). Applying machine learning EEG signal classification to emotion-related brain anticipatory activity. *F1000Research* 9:173. doi: 10.12688/f1000research.22202.2

Bilucaglia, M., Masi, R., Stanislao, G. D., Laureanti, R., Fici, A., Circi, R., et al. (2020). ESB: a low-cost EEG synchronization box. *HardwareX* 8:e00125. doi: 10.1016/j.ohx.2020.e00125

Blinowska, K. J., and Żygierewicz, J. (2021). Practical Biomedical Signal Analysis Using MATLAB®. Boca Raton, FL: CRC Press.

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Calvo, R. A., and D'Mello, S. (2010). Affect detection: an interdisciplinary review of models, methods, and their applications. *IEEE Transact. Affect. Comp.* 1, 18–37. doi: 10.1109/T-AFFC.2010.1

Can, Y. S., Mahesh, B., and André, E. (2023). Approaches, applications, and challenges in physiological emotion recognition—a tutorial overview. *Proc. IEEE* 111, 1287–1313. doi: 10.1109/JPROC.2023.3286445

Casado-Aranda, L.-A., Sánchez-Fernández, J., Bigne, E., and Smidts, A. (2023). The application of neuromarketing tools in communication research: a comprehensive review of trends. *Psychol. Market*. 40, 1737–1756. doi: 10.1002/mar.21832

Chang, C.-Y., Hsu, S.-H., Pion-Tonachini, L., and Jung, T.-P. (2018). "Evaluation of artifact subspace reconstruction for automatic EEG artifact removal," in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (Honolulu, HI), 1242–1245.

Chen, J., Ro, T., and Zhu, Z. (2022). Emotion recognition with audio, video, EEG, and EMG: a dataset and baseline approaches. *IEEE Access* 10, 13229–13242. doi: 10.1109/ACCESS.2022.3146729

Cherubino, P., Caratù, M., Modica, E., Rossi, D., Trettel, A., Maglione, A. G., et al. (2017). "Assessing cerebral and emotional activity during the purchase of fruit and vegetable products in the supermarkets," in *Neuroeconomic and Behavioral Aspects* of *Decision Making: Proceedings of the 2016 Computational Methods in Experimental Economics (CMEE) Conference* (Cham: Springer), 293–307.

Cherubino, P., Martinez-Levy, A. C., Caratu, M., Cartocci, G., Di Flumeri, G., Modica, E., et al. (2019). Consumer behaviour through the eyes of neurophysiological measures: state-of-the-art and future trends. *Comput. Intell. Neurosci.* 2019:1976847. doi: 10.1155/2019/1976847

Ciceri, A., Russo, V., Songa, G., Gabrielli, G., and Clement, J. (2020). A neuroscientific method for assessing effectiveness of digital vs. print ads: Using biometric techniques to measure cross-media ad experience and recall. *J. Advert. Res.* 60, 71–86. doi: 10.2501/JAR-2019-015

Cohen, J. (2013). Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge.

Dall'Olio, L., Curti, N., Remondini, D., Safi Harb, Y., Asselbergs, F. W., Castellani, G., et al. (2020). Prediction of vascular aging based on smartphone acquired ppg signals. *Sci. Rep.* 10:19756. doi: 10.1038/s41598-020-76816-6

Delorme, A., and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

D'Mello, S., Kappas, A., and Gratch, J. (2017). The affective computing approach to affect measurement. *Emot. Review* 10, 174–183. doi: 10.1177/1754073917696583

D'Mello, S., and Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. ACM Comp. Surv. 47, 1–36. doi: 10.1145/2682899

Dong, L., Li, F., Liu, Q., Wen, X., Lai, Y., Xu, P., et al. (2017). MATLAB toolboxes for reference electrode standardization technique (REST) of scalp EEG. *Front. Neurosci.* 11:601. doi: 10.3389/fnins.2017.00601

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G* power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychol. Bull.* 76:378. doi: 10.1037/h0031619

Funder, D. C., and Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* 2, 156–168. doi: 10.1177/2515245919847202

Gucciardi, A., Crotti, M., Itzhak, N. B., Mailleux, L., Ortibus, E., Michelucci, U., et al. (2022). "A new median filter application to deal with large windows of missing data in eye-gaze measurements," in *Proceedings of the International Workshop on Neurodevelopmental Impairments in Preterm Children - Computational Advancements* (Ljubljana), 1–17.

Gunes, H., and Pantic, M. (2010). Automatic, dimensional and continuous emotion recognition. *Int. J. Synth. Emot.* 1, 68–99. doi: 10.4018/jse.2010101605

Hakim, A., and Levy, D. J. (2018). A gateway to consumers' minds: achievements, caveats, and prospects of electroencephalography-based prediction in neuromarketing. *Wiley Interdiscipl. Rev. Cogn. Sci.* 10:e1485. doi: 10.1002/wcs.1485

Hernandez, J., Lovejoy, J., McDuff, D., Suh, J., O'Brien, T., Sethumadhavan, A., et al. (2021). "Guidelines for assessing and minimizing risks of emotion recognition applications," in 2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII) (Nara: IEEE), 1–8.

Imani, M., and Montazer, G. A. (2019). A survey of emotion recognition methods with emphasis on e-learning environments. *J. Netw. Comp. Appl.* 147:102423. doi: 10.1016/j.jnca.2019.102423

Kar, A. (2020). Mlgaze: Machine learning-based analysis of gaze error patterns in consumer eye tracking systems. *Vision* 4:25. doi: 10.3390/vision4020025

Karmarkar, U. R., and Plassmann, H. (2017). Consumer neuroscience: past, present, and future. Organ. Res. Methods 22, 174–195. doi: 10.1177/1094428117730598

Katsigiannis, S., and Ramzan, N. (2017). Dreamer: A database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices. *IEEE J. Biomed. Health Inf.* 22, 98–107. doi: 10.1109/JBHI.2017.2688239

Kleckner, I. R., Jones, R. M., Wilder-Smith, O., Wormwood, J. B., Akcakaya, M., Quigley, K. S., et al. (2018). Simple, transparent, and flexible automated quality assessment procedures for ambulatory electrodermal activity data. *IEEE Transact. Biomed. Eng.* 65, 1460–1467. doi: 10.1109/TBME.2017.27 58643

Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., et al. (2011). Deap: a database for emotion analysis; using physiological signals. *IEEE Transact. Affect. Comp.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15

Kurdi, B., Lozano, S., and Banaji, M. R. (2017). Introducing the open affective standardized image set (oasis). *Behav. Res. Methods* 49, 457–470. doi: 10.3758/s13428-016-0715-3

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33:159. doi: 10.2307/2529310

Lang, P., Bradley, M., and Cuthbert, B. (2008). *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*. Technical Report, University of Florida, Gainesville, FL, United States.

Larradet, F., Niewiadomski, R., Barresi, G., Caldwell, D. G., and Mattos, L. S. (2020). Toward emotion recognition from physiological signals in the wild: approaching the methodological issues in real-life data collection. *Front. Psychol.* 11:1111. doi: 10.3389/fpsyg.2020.01111

Lee, N., Brandes, L., Chamberlain, L., and Senior, C. (2017). This is your brain on neuromarketing: reflections on a decade of research. *J. Market. Manag.* 33, 878–892. doi: 10.1080/0267257X.2017.1327249

Lerner, J. S., Li, Y., Valdesolo, P., and Kassam, K. S. (2015). Emotion and decision making. Annu. Rev. Psychol. 66, 799-823. doi: 10.1146/annurev-psych-010213-115043

Lim, W. M. (2018). What will business-to-business marketers learn from neuromarketing? Insights for business marketing practice. J. Bus. Bus. Market. 25, 251–259. doi: 10.1080/1051712X.2018.1488915

Liu, B., Zhao, Q.-C., Ren, Y.-Y., Wang, Q.-J., and Zheng, X.-L. (2018). An elaborate algorithm for automatic processing of eye movement data and identifying fixations in eye-tracking experiments. *Adv. Mech. Eng.* 10:168781401877367. doi: 10.1177/1687814018773678

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., et al. (2019). Jasp: graphical statistical software for common statistical designs. *J. Stat. Softw.* 88, 1–17. doi: 10.18637/jss.v088.i02

Luck, S. J. (2022). Applied Event-Related Potential Data Analysis. LibreTexts.

Mancini, M., Cherubino, P., Martinez, A., Vozzi, A., Menicocci, S., Ferrara, S., et al. (2023). What is behind in-stream advertising on youtube? A remote neuromarketing study employing eye-tracking and facial coding techniques. *Brain Sci.* 13:1481. doi: 10.3390/brainsci13101481

Merletti, R., and Cerone, G. (2020). Tutorial. Surface emg detection, conditioning and pre-processing: best practices. J. Electromyogr. Kinesiol. 54:102440. doi: 10.1016/j.jelekin.2020.102440

Miranda-Correa, J. A., Abadi, M. K., Sebe, N., and Patras, I. (2018). Amigos: A dataset for affect, personality and mood research on individuals and groups. *IEEE Transact. Affect. Comp.* 12, 479–493. doi: 10.1109/TAFFC.2018.2884461

Nathan, M. J., and Pinal, G. D. (2017). The future of cognitive neuroscience? Reverse inference in focus. *Philos. Compass* 12:e12427. doi: 10.1111/phc3.12427

Nemorin, S. (2016). Neuromarketing and the "poor in world" consumer: how the animalization of thinking underpins contemporary market research discourses. *Consumpt. Mark. Cult.* 20, 59–80. doi: 10.1080/10253866.2016.1160897

Nuwer, M. R. (2018). 10-10 electrode system for EEG recording. *Clin. Neurophysiol.* 129:1103. doi: 10.1016/j.clinph.2018.01.065

Pant, S., Yang, H.-J., Lim, E., Kim, S.-H., and Yoo, S.-B. (2023). PhyMER: physiological dataset for multimodal emotion recognition with personality as a context. *IEEE Access* 11, 107638–107656. doi: 10.1109/ACCESS.2023.3320053

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* 9, 148–158. doi: 10.1038/nrn2317

Picard, R. W. (1995). Affective Computing. Technical Report 321. Cambridge, MA: M.I.T Media Laboratory Perceptual Computing Section.

Picard, R. W. (1997). Affective Computing. Cambridge, MA: MIT Press.

Pion-Tonachini, L., Kreutz-Delgado, K., and Makeig, S. (2019). Iclabel: an automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* 198, 181–197. doi: 10.1016/j.neuroimage.2019.05.026

Plassmann, H., Ramsøy, T. Z., and Milosavljevic, M. (2012). Branding the brain: a critical review and outlook. *J. Consum. Psychol.* 22, 18–36. doi: 10.1016/j.jcps.2011.11.010

Plassmann, H., Venkatraman, V., Huettel, S., and Yoon, C. (2015). Consumer neuroscience: applications, challenges, and possible solutions. *J. Market. Res.* 52, 427-435. doi: 10.1509/jmr.14.0048

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63. doi: 10.1016/j.tics.2005.12.004

Ramsøy, T. Z. (2019). Building a foundation for neuromarketing and consumer neuroscience research. J. Advert. Res. 59, 281–294. doi: 10.2501/JAR-2019-034

Rangayyan, R. M., and Krishnan, S. (2024). *Biomedical Signal Analysis*. Hoboken, NJ: John Wiley & Sons.

Rawnaque, F. S., Rahman, K. M., Anwar, S. F., Vaidyanathan, R., Chau, T., Sarker, F., et al. (2020). Technological advancements and opportunities in neuromarketing: a systematic review. *Brain Inf.* 7, 1–19. doi: 10.1186/s40708-020-00109-x

Russo, V., Bilucaglia, M., Casiraghi, C., Chiarelli, S., Columbano, M., Fici, A., et al. (2023). Neuroselling: applying neuroscience to selling for a new business perspective. An analysis on teleshopping advertising. *Front. Psychol.* 14:1238879. doi: 10.3389/fpsyg.2023.1238879

Russo, V., Bilucaglia, M., Circi, R., Bellati, M., Valesi, R., Laureanti, R., et al. (2022). The role of the emotional sequence in the communication of the territorial cheeses: a neuromarketing approach. *Foods* 11:2349. doi: 10.3390/foods11 152349

Russo, V., Milani Marin, L. E., Fici, A., Bilucaglia, M., Circi, R., Rivetti, F., et al. (2021). Strategic communication and neuromarketing in the fisheries sector: generating ideas from the territory. *Front. Commun.* 6:49. doi:10.3389/fcomm.2021.659484

Russo, V., Valesi, R., Gallo, A., Laureanti, R., and Zito, M. (2020). "The theater of the mind": the effect of radio exposure on TV advertising. *Soc. Sci.* 9:123. doi: 10.3390/socsci9070123

Saganowski, S., Komoszyńska, J., Behnke, M., Perz, B., Kunc, D., Klich, B., et al. (2022). Emognition dataset: emotion recognition with self-reports, facial expressions, and physiology using wearables. *Scientific Data* 9(1). doi: 10.1038/s41597-022-0 1262-0

Schmidt, P., Reiss, A., D⁵urichen, R., and Laerhoven, K. V. (2019). Wearable-based affect recognition-review. *Sensors*. 19, 4079. doi: 10.3390/s19194079

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., et al. (2018). A review of emotion recognition using physiological signals. *Sensors* 18:2074. doi: 10.3390/s18072074

Siddiqui, M. F. H., Dhakal, P., Yang, X., and Javaid, A. Y. (2022). A survey on databases for multimodal emotion recognition and an introduction to the VIRI (visible and InfraRed image) database. *Multim. Technol. Interact.* 6:47. doi: 10.3390/mti6060047

Sinha, S. R., Sullivan, L., Sabau, D., San-Juan, D., Dombrowski, K. E., Halford, J. J., et al. (2016). American clinical neurophysiology society guideline 1: minimum technical requirements for performing clinical electroencephalography. J. Clin. Neurophysiol. 33, 303–307. doi: 10.1097/WNP.000000000000308

Smith, A. (2002). Kijken in Het Brein. Over de Mogelijkheden van Neuromarketing. Technical Report 321. Rotterdam: Erasmus Research Institute of Management. Soleymani, M., Lichtenauer, J., Pun, T., and Pantic, M. (2011). A multimodal database for affect recognition and implicit tagging. *IEEE Transact. Affect. Comp.* 3, 42–55. doi: 10.1109/T-AFFC.2011.25

Song, T., Zheng, W., Lu, C., Zong, Y., Zhang, X., and Cui, Z. (2019). Mped: a multimodal physiological emotion database for discrete emotion recognition. *IEEE Access* 7, 12177–12191. doi: 10.1109/ACCESS.2019.2891579

Soodan, V., and Pandey, A. C. (2016). Influence of emotions on consumer buying behavior. J. Entrepreneurship Bus. Econ. 4, 163–181. Available online at: https://scientificia.com/index.php/JEBE/article/view/48

Stasi, A., Songa, G., Mauri, M., Ciceri, A., Diotallevi, F., Nardone, G., et al. (2018). Neuromarketing empirical approaches and food choice: a systematic review. *Food Res. Int.* 108, 650–664. doi: 10.1016/j.foodres.2017.11.049

Subramanian, R., Wache, J., Abadi, M. K., Vieriu, R. L., Winkler, S., and Sebe, N. (2016). Ascertain: emotion and personality recognition using commercial sensors. *IEEE Transact. Affect. Comp.* 9, 147–160. doi: 10.1109/TAFFC.2016.2625250

Tyng, C. M., Amin, H. U., Saad, M. N., and Malik, A. S. (2017). The influences of emotion on learning and memory. *Front. Psychol.* 8:1454. doi: 10.3389/fpsyg.2017.01454

Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., et al. (2022). A systematic review on affective computing: emotion models, databases, and recent advances. *Inf. Fus.* 83–84, 19–52. doi: 10.1016/j.inffus.2022.03.009

Yoon, C., Gutchess, A. H., Feinberg, F., and Polk, T. A. (2006). A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *J. Cons. Res.* 33, 31–40. doi: 10.1086/504132

Zhang, J., Yin, Z., Chen, P., and Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: a tutorial and review. *Inf. Fus.* 59, 103–126. doi: 10.1016/j.inffus.2020.01.011

Zhang, L., Walter, S., Ma, X., Werner, P., Al-Hamadi, A., Traue, H. C., et al. (2016). ""biovid emo db": A multimodal database for emotion analyses validated by subjective ratings," in 2016 IEEE Symposium Series on Computational Intelligence (SSCI) (Athens: IEEE), 1–6.

Zhao, S., Jia, G., Yang, J., Ding, G., and Keutzer, K. (2021). Emotion recognition from multiple modalities: Fundamentals and methodologies. *IEEE Signal Process. Mag.* 38, 59–73. doi: 10.1109/MSP.2021.3106895

Zheng, W.-L., Zhu, J.-Y., and Lu, B.-L. (2017). Identifying stable patterns over time for emotion recognition from eeg. *IEEE Transact. Affect. Comp.* 10, 417–429. doi: 10.1109/TAFFC.2017.2712143

Zito, M., Fici, A., Bilucaglia, M., Ambrogetti, F. S., and Russo, V. (2021). Assessing the emotional response in social communication: the role of neuromarketing. *Front. Psychol.* 12:625570. doi: 10.3389/fpsyg.2021.625570