



OPEN

DATA DESCRIPTOR

WEMAC: Women and Emotion Multi-modal Affective Computing dataset

Jose A. Miranda Calero^{1,2,6}✉, Laura Gutiérrez-Martín^{3,6}, Esther Rituerto-González^{2,4,6}✉, Elena Romero-Perales^{2,3}, Jose M. Lanza-Gutiérrez⁵, Carmen Peláez-Moreno^{2,4} & Celia López-Ongil^{2,3}

WEMAC is a unique open multi-modal dataset that comprises physiological, speech, and self-reported emotional data records of 100 women, targeting Gender-based Violence detection. Emotions were elicited through visualizing a validated video set using an immersive virtual reality headset. The physiological signals captured during the experiment include blood volume pulse, galvanic skin response, and skin temperature. The speech was acquired right after the stimuli visualization to capture the final traces of the perceived emotion. Subjects were asked to annotate among 12 categorical emotions, several dimensional emotions with a modified version of the Self-Assessment Manikin, and liking and familiarity labels. The technical validation proves that all the targeted categorical emotions show a strong statistically significant positive correlation with their corresponding reported ones. That means that the videos elicit the desired emotions in the users in most cases. Specifically, a negative correlation is found when comparing fear and not-fear emotions, indicating that this is a well-portrayed emotional dimension, a specific, though not exclusive, purpose of WEMAC towards detecting gender violence

Background & Summary

Gender-based Violence (GBV) is a violation of human rights and fundamental liberties as declared by the United Nations in 1993¹. This declaration states that GBV is any act of physical, sexual, or psychological violence directed toward the female gender. A datum that helps ponder its impact on society shows that more than 27% of ever-partnered women aged between 15 and 49 experienced physical or sexual violence by intimate partners from 2000 to 2018². Another worrying statistic is that, in 2020, approximately 47, 000 women and girls were killed worldwide by their intimate partners or other family members, meaning a woman or girl is killed by someone in her own family every 11 minutes³. On this basis, it can be understood that GBV is an urgent problem worldwide that should be addressed both by society and science. From a sociological point of view, particular emphasis on education is essential to eradicate, combat, and prevent GBV, but it requires several generations to produce a change in society. In the meantime, technology can be fundamental for helping prevent and combat GBV⁴ and empowering women. With this focus, the multidisciplinary team UC3M4Safety was created in 2017 to propose Bindi, an inconspicuous autonomous system powered by artificial intelligence and deployed under the Internet of Things (IoT) paradigm. The goal of Bindi is to automatically report when a woman is in a GBV-related risk situation to trigger a protection protocol⁵. This risk situation identification is performed by detecting fear-related emotions in the user through a multi-modal intelligent engine. This engine is fed by inconspicuous physiological and audio sensors embodied into a pair of wearable devices: a bracelet and a pendant.

Despite the wealth of research into affective states or emotion detection using auditory or physiological cues for diverse purposes^{6–8}, there remains a distinct lack of focus on detecting fear-related states that could denote risky situations, as in GBV scenarios. This gap presents an urgent opportunity to craft emotion recognition

¹Embedded Systems Laboratory, Ecole Polytechnique Fédérale de Lausanne (EPFL), Vaud, Switzerland. ²Instituto de Estudios de Género, Universidad Carlos III de Madrid, Calle Madrid, 126, 28903, Getafe (Madrid), España.

³Departamento de Tecnología Electrónica, Universidad Carlos III de Madrid, Avenida de la Universidad, 30, 28911, Leganés (Madrid), España. ⁴Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, Avenida de la Universidad, 30, 28911, Leganés (Madrid), España. ⁵Departamento de Ciencias de la Computación, Universidad de Alcalá, Escuela Politécnica Superior, Madrid, España. ⁶These authors contributed equally: Jose A. Miranda Calero, Laura Gutiérrez-Martín, Esther Rituerto-González. ✉e-mail: jose.mirandacalero@epfl.ch; erituert@ing.uc3m.es

Database	Datasets	Conditions	Participants
UC3M4Safety Database ²²	Audiovisual Stimuli: Videos ²³	Online crowd-sourcing	General public and expert judges
	Audiovisual Stimuli: Emotional Ratings ²⁴		
	WEMAC: Biopsychosocial Questionnaire ²⁵	Laboratory	Women volunteers
	WEMAC: Physiological Signals ²⁶		
	WEMAC: Audio Features ²⁷		
	WEMAC: Self-reported Emotional Annotations ²⁸		

Table 1. Hierarchy and subdivisions of the UC3M4Safety Database datasets.

solutions tailored specifically for GBV cases. Moreover, leveraging multimodal data fusion of different modalities offers substantial advantages over unimodal approaches, a fact that's increasingly recognized and on the rise in current research^{9–14}. However, the majority of emotion recognition systems currently do not integrate data fusion of physiological and auditory data, nor do they focus on the needs of vulnerable groups like GBV victims. Notably, these two specific modalities can be seamlessly embodied and integrated into resource constrained wearable devices, offering considerably more inconspicuous solutions rather than using other signals such as electroencephalogram (EEG).

One of the main shortcomings of training this specific *fear* detection system is the lack of adequate datasets. Over the last decades, several datasets were published providing emotional labels together with auditory and physiological variables, such as DEAP¹⁵, MAHNOB¹⁶, WESAD¹⁷, AMIGOS¹⁸, FAU, Reg, and Ulm TSST Corpora¹⁹, and BioSpeech²⁰. However, the target of these datasets is the classification of a generic set of emotions instead of focusing on *fear* detection. This strategy makes it hard to obtain a robust model due to the lack of *fear* samples. Moreover, gender perspective was not considered either, despite stimuli interpretation being strongly affected by gender²¹. Instead, a sufficient number of women volunteers and a balanced *fear* and *non-fear* set of emotions should be considered to fit the target users of the GBV application.

To fill this gap, UC3M4Safety generated the UC3M4Safety Database²², which includes the Women and Emotion Multi-modal Affective Computing (WEMAC) and audiovisual stimuli datasets, as listed in Table 1. This paper presents WEMAC, a collection of physiological and audio data captured in a virtual reality set-up where women volunteers are exposed to a subset of audiovisual stimuli previously rated by experts judges and an online crowd-sourcing procedure^{23,24}. For each volunteer, WEMAC includes the answers to an initial questionnaire²⁵, physiological signals²⁶, features extracted from the volunteer's speech data—to ensure privacy and confidentiality²⁷, and self-reported emotional labels²⁸.

WEMAC offers several advantages over other state-of-the-art datasets: 1) the use of immersive technology (virtual reality) to elicit emotions, which offers a high degree of correlation between the research conditions and the emotional phenomenon under study; 2) a high number of volunteers (100), higher than any other public dataset of this type to the best of our knowledge; 3) a specifically designed modification of the labeling methodology to consider gender perspective by changing the design of the original Self-Assessment Manikins (SAMs)²⁹; 4) the implementation of an online recovery process to ensure a physiological stabilization between stimuli, and 5) the usage of different sensory systems to provide a heterogeneous set-up approach.

WEMAC is intended, but not limited, to address research questions related to 1) affective computing using multi-modal information, 2) the design of solutions to the very challenging problem of GBV, 3) the understanding of subjective and self-reported emotional labeling, 4) the analysis of physiological signal differences between different emotions, 5) the impact of using 2D and 3D audio-visual stimuli, and 6) *fear* classification in women.

Methods

Ethics statement. The experimentation was approved by and performed following the guidelines and regulations of the Ethics in Research Committee of University Carlos III Madrid. The approval was granted considering the circumstances of the research project entitled *Integral protection of gender-based violence victims using multimodal affective computing*. Note that the original Spanish title is *Protección Integral de las víctimas de violencia de género mediante Computación Afectiva multimodal* (EMPATIA-CM) with reference Y2018/TCS-5046 and funded by the programme *Proyectos Sinérgicos I+D (Comunidad de Madrid, Consejería de Ciencia, Universidades e Innovación)*, whose principal investigator is Dr. Celia López-Ongil.

The submission to the Ethical Committee covered essential topics for the development of the experiments. Among others, the adequacy of the volunteers' informed consent, the research goals and plans, the data management and de-identification procedures, and the compliance with the European General Data Protection Regulation (GDPR). The aforementioned written informed consent asserts that the volunteers 1) were aware of the research, its objectives, purposes, and how their data was planned to be used; 2) were informed regarding the experimental procedure, the possibility of refusing to participate in the research at any point, and the right to request data erasure; 3) could ask any question during the experiments. Moreover, they granted permission for processing their personal data to the extent necessary for the implementation of the research project, including sharing with other researchers their physiological recording and speech features, as well as the initial questionnaires and self-reported annotations.

Participants. The volunteers were recruited by using different communication channels, including social networks (such as Facebook, Twitter, or Instagram), Women's Associations, and the collaboration of the municipality

of Leganés and Getafe (Madrid, Spain), resulting in an initial selection of 144 Spanish-speaking women volunteers. This initial set was reduced due to virtual reality sickness^{30,31}, a phenomenon that appears when using head-mounted displays, and volunteers' discomfort when facing some sensitive stimuli, both involving interruption of the experiment. Besides some technical problems arose during the recording. The final set includes 100 women volunteers aged between 20 and 77 (with an average of 39.92 and a standard deviation of 14.26). To cover a wide age range ensuring equal distribution across all ages, or at least across decades, the recruitment process requested a balanced number of volunteers in five age groups defined as G1 (18–24) for which we recruited 22 volunteers, G2 (25–34) with 18 volunteers and G3 (35–44), G4 (45–54), and G5 (≥ 55) that include 20 volunteers each. Other distributions could be generated by joining groups, for instance, to analyze the influence of age on any of the data collected.

Stimuli. The selection of audiovisual clips to record WEMAC is the result of a thorough previous study done to collect a high-quality set of audiovisual stimuli able to trigger emotions under a controlled scenario, which is detailed in³². To this end, 370 samples with emotional content were initially collected on the Internet from commercial films, series, documentaries, short films, commercials, and video clips. The set was labeled with the advice of a panel of experts seeking to elicit the 12 basic discrete emotions obtained in the previous work: *joy, sadness, surprise, contempt, hope, fear, attraction, disgust, tenderness, anger, calm, and tedium*. After that, the team discarded those clips over two minutes long, those needing context to be understood, or that elicited more than one target emotion. Afterward, the resulting 162 clips were surveyed in an online crowd-sourcing poll to be labeled with discrete emotion categories after its visualization, getting 1,520 volunteer annotators (929 women and 591 men). A further selection was performed by considering two conditions: 1) to ensure that each of the selected video clips represented a predominant emotion, with at least 50% agreement among the evaluators; 2) to verify the exclusivity of the assigned emotion by verifying that all other possible emotions had a maximum agreement of only 30%. Note that the latter threshold was empirically determined. After this selection, none of the videos previously selected for *attraction, contempt, hope, and tedium* had greater than 50% agreement among the responses obtained, so the final set of videos covers a list of 8 target emotions *joy, sadness, surprise, fear, attraction, disgust, tenderness, anger, and calm*. However, it is important to acknowledge that participants may experience any of the 12 basic emotions while facing the stimuli, not just the eight specific emotions identified in the selection. Finally, some videos were discarded to obtain an even distribution between *fear* and *non-fear* emotions to facilitate the generalization of the machine learning model and prevent an unbalanced database. It resulted in the Audiovisual Stimuli dataset²³ with 42 clips and a distribution of 44.44% for *fear* and 55.55% for *non-fear*²⁴.

Thus, out of the whole Audiovisual Stimuli dataset in WEMAC, only two batches were generated due to the difficulties and time-consuming process of recruiting volunteers. Besides, a significant emphasis was placed on ensuring that an adequate and balanced number of participants were enrolled in each batch, with appropriate representation across varying age groups. For each batch a total number of 14 videos were selected to make the acquisition process feasible, and to ensure that the experimental procedure lasted 1 to 1.5 hours per volunteer. The selection criterion for the videos in these two batches is based on three premises: the emotional highest discrete labeling agreement, targeting for an adequate laboratory experiment duration, and a balanced distribution of *fear* and *non-fear* categories within the four quadrants in the Valence-Arousal space³³. Table 2 shows details about the 14 selected videos for each batch in WEMAC, including the stimuli identification (Stimuli ID, with the same notation as followed in the Audiovisual Stimuli dataset²³), the visualization order in the experimentation, the emotion label reported in the online crowd-sourcing study, the video duration, and the visualization format. A limitation of the WEMAC dataset, is a possible ordering effect during the acquisition protocol's design. From our point of view, it is necessary to count on a very and considerably larger number of volunteers to completely and effectively neutralize the order effect through random video clip selection, a joint requirement that is not met by any similar existing database. Moreover, to ameliorate such effect in our case, the dataset ensured equally distributed *fear* and *non-fear* emotion videos, with *fear*-emotion videos interleaved across both batches.

Measures. Volunteers' annotations are collected in two instants: prior to the experiment and during the experimentation. Before the experiment, each volunteer is provided with informed consent, a personal data form, and a general questionnaire to supply additional information related to cognition, appraisal, attention, personality traits, gender, and age. In this regard, the general questionnaire collects age group, recent physical activity or medication that can alter the physiological response of the participant, self-identified emotional burdens due to work, economic and personal situation, and mood biases (fears, phobias, and previous traumatic experiences).

During the experimental protocol, in addition to physiological information, the following self-assessment annotations are obtained after each visualized stimulus:

- Speech-based labeling: two questions regarding each video stimulus are asked immediately after its visualization. Their goal is to make the user relive the emotions felt during viewing. Some examples are "What did you feel during the visualization of the video?" or "Could you describe what happened in the video using your own words?". The answer is stored as an audio signal. Note that both questions and answers are in Spanish.
- Dimensional emotion ratings (Valence, Arousal, and Dominance)³⁴: annotated using a 9-point Likert scale supported by modified SAMs. Details about the redesign process of the SAMs can be found in²⁹.
- Familiarity with the emotion felt, the situation displayed in the clip, and the specific clip: annotated in three different questions. The two first consider a 9-point Likert scale, whereas the last one considers a binary yes-no option.
- Liking of the video: annotated through a binary yes-no question.
- A discrete emotion out of a total of 12 (*joy, sadness, surprise, contempt, hope, fear, attraction, disgust, tenderness, anger, calm, and tedium*)³².

Stimuli ID	Visualization order	Emotion label	Duration	Format	Batch
V01	1	Joy	1'26"	2D	1
V15	2	Fear	1'20"	3D	1
V36	3	Sadness	1'59"	2D	1
V08	4	Anger	1'03"	3D	1
V28	5	Fear	1'35"	2D	1
V40	6	Calm	1'	3D	1
V09	7	Anger	1'	2D	1
V19	8	Fear	23"	2D	1
V52	9	Disgust	40"	2D	1
V16	10	Fear	2'	3D	1
V02	11	Joy	1'41"	2D	1
V27	12	Fear	1'20"	2D	1
V37	13	Gratitude	1'40"	2D	1
V24	14	Fear	1'27"	2D	1
V22	1	Fear	1'52"	2D	2
V04	2	Joy	1'28"	2D	2
V11	3	Fear	46"	2D	2
V34	4	Sadness	45"	2D	2
V13	5	Fear	1'33"	3D	2
V41	6	Calm	1'	2D	2
V10	7	Anger	1'59"	2D	2
V25	8	Fear	1'14"	2D	2
V33	9	Disgust	1'36"	2D	2
V14	10	Fear	2'	3D	2
V07	11	Surprise	1'41"	2D	2
V26	12	Fear	1'06"	2D	2
V77	13	Gratitude	1'30"	2D	2
V12	14	Fear	1'59"	3D	2

Table 2. List of selected audio-visual stimuli used within the WEMAC Dataset.

Apparatus. The equipment employed to capture the physiological and speech data is as follows:

- The BioSignalPlux research toolkit system. It is a commonly used device to acquire different physiological signals in the literature^{35–38}. More specifically, we capture finger Blood Volume Pulse (BVP), ventral wrist Galvanic Skin Response (GSR), forearm Skin Temperature (SKT), trapezoidal Electromyography (EMG), chest respiration (RESP), and inertial wrist movement through an accelerometer. The sampling frequency for all physiological signals is 200 Hz.
- The Oculus Rift® S VR Headset. Its embedded microphone captures the speech signal produced during the speech-based labeling at a sampling rate of 48 kHz mono and a depth of 16 bits. This device also guides volunteers through the study. To this end, an interactive virtual reality environment developed in Unity® software presents stimuli and collects self-assessments. Note that upon request, we can send the full source code employed in the Unity® environment. The presentation of audiovisual clips varies depending on their type. 2D clips utilize a wide screen with black surroundings while 3D clips use a 360-degree spherical panorama. More details about this virtual environment can be found in³⁹.
- Additionally, two in-house sensory systems are employed. On the one hand, the Bindi's bracelet⁵ measures dorsal wrist BVP, ventral wrist GSR, and forearm SKT. The hardware and software particularities of this device are detailed in^{40–42}. Specifically, the MAX30101 and the MAX30208 integrated circuits are employed for the BVP and SKT measurements, respectively, offering high-sensitivity and clinical-grade measurements. For the GSR, the analog front-end is described in⁴² and the signal is acquired through the Analog-to-Digital Converter of the nRF52840 System-on-Chip. On the other hand, a GSR sensor to be integrated into the next version of the Bindi bracelet is used. Its hardware and software particularities are detailed in⁴³. The sampling frequency for all physiological signals is 200 Hz.

Note that the previously mentioned BioSignalPlux toolkit is employed as a golden standard to analyze the performance of the two in-house sensory systems due to their experimental nature. Driven by this latter fact, the physiological signals uploaded and published into WEMAC belong to the BioSignalPlux toolkit. Thus, the in-house sensory system signals are not publicly available, being subjected to an additional future dataset within the UC3M4Safety database after a complete end-to-end (i.e. signal integrity and emotion recognition classification vs golden standard) validation is performed.

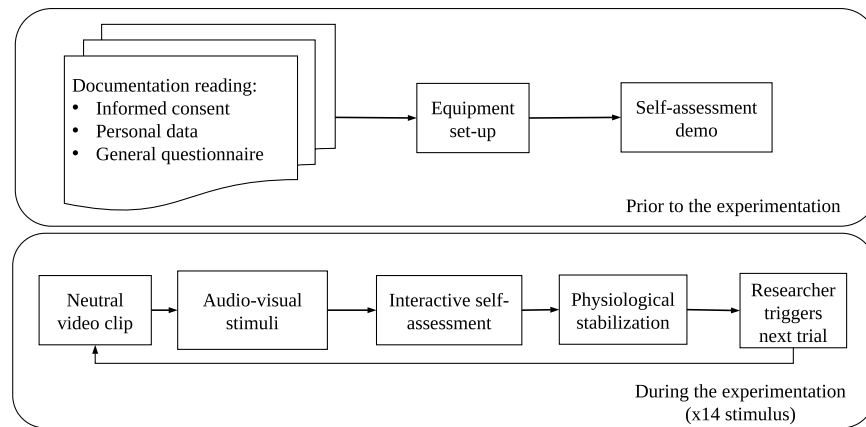


Fig. 1 Experimental methodology followed during the development of the WEMAC dataset. Prior to and during the experimentation for each volunteer.

The synchronization of all the different sensors acquisition together with the experiment stages is performed using a laptop (MSI GE75 Raider 8SE-034ES) running a Unity® framework-based program. On the one hand, the BiosignalPlux device connection is configured using the OpenSignals (r)evolution software, and its TCP/IP module is used to facilitate the data exchange between this platform and the Unity® framework. On the other hand, the additional in-house sensory systems are wirelessly connected to the laptop using Bluetooth Low Energy dongles. The information storage is divided by scenes and marked individually with a timestamp set by the environment employed.

Procedure. The study was conducted between October 2020 and April 2021. It took place in the Electronics Technology Department at the School of Engineering of the Universidad Carlos III de Madrid, Spain, and in the Gender Studies Institute at the same University. The experimental methodology designed to be applied for each volunteer is schematized in Figure 1. During this experiment, at least one researcher and one psychologist remained in the room at the disposal of the participants in case they needed any help.

Upon arrival, participants were informed about the experimental procedure. Then, they signed the informed consent, filled out the personal data form, and answered the general questionnaire. Next, participants listened to instructions regarding the experiment. The participants were requested to avoid unnecessary actions or movements during the experiment (e.g., turning the wrist). They were also informed that they could skip any clip or quit the experiment at any time. Once the procedure was clear to the participants, the sensors were set up, as well as the Virtual Reality Headset. Next, the participants followed a tutorial to get used to the headset, joystick, interactive screens, and the particularities of the different annotations. The main part of the experiment consisted of fourteen iterations (one per stimulus) of i) a neutral clip sampled from⁴⁴, ii) the visualization of the emotional film clip, iii) different interactive screens of self-assessment annotations, and iv) 3D recovery landscape scenes.

The full experiment, including documentation reading, equipment set-up, tutorial, and visualization of the clips, lasted around 80 minutes per participant, depending on the time spent on the questionnaires.

Data processing and cleaning. The whole set of physiological signals captured for the 100 volunteers has a total duration of 34:51:50 (HH:MM:SS), whereas the audio signals altogether last 12:19:25 since they are only saved during the speech-based annotation.

Physiological Signals. The signals being released and publicly available in WEMAC are the ones acquired by the BioSignalPlux research toolkit. Specifically, the raw and filtered BVP, GSR, and SKT signals captured during every video visualization are provided. The preprocessing is as follows:

1. Two main filters are applied to the BVP signal due to the noise problems observed. On the one hand, high-frequency noise is filtered out using a direct-form low-pass Finite Impulse Response (FIR) filter with 6dB at 3.5 Hz. Note that a Hamming window is used during the filter design process to minimize the first side lobe properly. On the other hand, the residual baseline-wander or low-frequency drift effect presented in the signal is removed using a forward-backwards low-pass Butterworth Infinite Impulse Response (IIR) filtering stage. Specifically, the forward-backwards technique handles the non-linear phase of such filters.
2. For the GSR and SKT signals, a basic FIR filtering with 2 Hz cut-off frequency is applied. After that, this filtered output is downsampled to 10 Hz and also processed with both a moving average and a moving median filter. The former used a 1s long window and helped to reduce the high noise residual after the initial FIR, whereas the latter employed a 0.5s window and dealt with the rapid transients.

Speech Signals. The speech signals recorded have a duration ranging from 20s to 60s and mostly contain speech. Since the release of the raw speech signals is not possible due to ethics and privacy issues, see regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural

persons concerning the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (*General Data Protection Regulation*), we preprocessed the speech signals and extracted both low and high-level features so that the research community can analyze and work with them.

The preprocessing is as follows. A low pass filter at 8 kHz is applied, where most of the energy is concentrated. A high pass filter with a cut-off frequency of 50 Hz is also applied to remove the electrical power line interference. Afterward, the audio amplitude is normalized per participant to fit the range $[-1, 1]$. That means the normalization per participant was performed using all her audio signals. Next, downsampling at 16 kHz is performed with the *librosa* Python library⁴⁵ to facilitate the handling of the signals. Finally, the signals are padded with zeros to fill in the incomplete last second.

Different Python toolkits are used to extract information from the preprocessed signal at a window size of 1 s without overlapping. We follow a similar approach to the one followed in the MuSe Challenge 2021⁴⁶ for the feature and embedding extraction of the audio signals. That is:

1. *librosa*^{45,47}: 19 features are extracted by the *librosa* Python toolkit (13 Mel-Frequency Cepstral Coefficients, Root Mean Square or Energy, Zero Crossing Rate, Spectral Centroid, Spectral Roll-off, Spectral Flatness, and Pitch) at a window size of 20 ms and a hop size of 10 ms. Also, the mean and standard deviation for each of the 19 features are computed for every 1 s window. As a result, 38 speech features are computed.
2. *eGeMAPS*⁴⁸: 88 Low-Level Descriptors (LLDs) related to speech and audio are extracted through the *openSMILE* Python toolkit⁴⁹ on its default configuration, i.e., F0, harmonic features, HNR, jitter, and shimmer are computed with a 60 ms window. Loudness, spectral slope, formants, harmonics, Hammarberg Index, and Alpha ratio are computed with a 25 ms window. They are all further averaged in 1 s windows.
3. *ComParE*: 6, 373 features used in the ComParE 2016 challenge⁵⁰ are extracted by using the *openSMILE* Python toolkit. The default window and step sizes are used as in⁵¹, i.e., F0, jitter, and shimmer are computed with a window size of 60 ms and a step of 10 ms. All other LLDs are computed with a window size of 20 ms and a step of 10 ms. Then all features are further averaged in 1 s windows.
4. *DeepSpectrum*⁵²: This toolkit is used for the extraction of audio embeddings based on different deep neural network architectures trained with ImageNet⁵³. Specifically, two different configurations are considered, and therefore two embeddings sets are extracted, i.e., ResNet50 and the output of its last Average Pooling layer (*avg_pool*), resulting in 2048-dimensional embeddings, and VGG-19 and its last Fully Connected layer (*fc2*), resulting in 4096-dimensional embeddings.
5. *VGGish*: The 128-dimensional embeddings from the output layer of the VGG-19 network trained for AudioSet⁵⁴ are also included.
6. *PASE+*: The 256-dimensional features from the PASE+ (*Problem Agnostic Speech Encoder+*)⁵⁵ encoder network, used as speech feature extractor, are also provided.

Data Records

The collected data in WEMAC (physiological signals, audio features, self-reported annotations, and general questionnaire answers) are available in *e-cienciaDatos* portal²², a research data repository held by Madroño Consortium (composed of Universities in the Madrid-Spain region and member of the Harvard Dataverse Network, accessible for major publishers). The information is split into four sections according to their typology, where each set contains a 'README.txt' file that explains each specific content.

1. *Bio-psycho-social questionnaire and informed consent*²⁵: a tabular data CSV (Comma-Separated Values) file with 100 rows corresponding to each of the volunteers and 14 columns corresponding to the information completed in the biopsychosocial questionnaire at the beginning of the experiment. This information includes: (1) volunteer identifier; (2) the age group in which she was included; (3) if she had signed the informed consent; (4) the number of children she had if any; (5) if she had had energy drinks before the experiment; (6) if she was taking medicines; (7) if she usually did physical exercise; (8) if she considered herself fearful; (9) if she had any deepest fear; (10) if she was in a stressful period; (11) if she had experienced any traumatic experiences; (12) if social interaction made her feel anxious; (13) if she considered having a tendency to be worried; (14) if she had ever felt fear of aggression; and (15) if she had ever felt a threat to her sexual integrity.
2. *Self-reported annotations*²⁸: They contain the emotional labeling reported by the participants after watching each of the 14 videos in the experiment. The data are stored in one CSV file, that contains 14 columns and 1,400 rows (100 volunteers \times 14 clips). Regarding the columns, 5 of them refer to information about the volunteer (identifier and age group) and the video (batch, display position, and video code number), whereas 9 of them refers to the self-assessment provided (arousal, valence, dominance, liking, reported and target discrete emotions, and familiarity scores for the emotion, the situation, and the clip).
3. *Physiological signals*²⁶: BVP, GSR, and SKT physiological signals captured during the experimentation by the BioSignalPlux research toolkit are provided in a binary MATLAB[®] file (.mat). It contains a cell array with 100 rows (one per volunteer) and 14 columns (one per video). Each cell contains four fields: volunteer identifier, clip or trial identifier, filtering indicator, and an inner cell array (with the physiological data associated with that specific clip and volunteer).
4. *Speech features*²⁷: The speech features contain 6 folders: *librosa*, *eGeMAPS*, *Compare*, *Deepspectrum-Resnet50*, *Deepspectrum-VGG19*, and *VGGish*. They correspond to the six feature sets described in the Data Processing and Cleaning section for Speech Signals. Each folder includes a CSV file per audiovisual stimuli and volunteer. Each CSV has as many columns as the number of features calculated, where an additional first column is also included referring to the timestamp (in seconds). The number of rows fits with the number of seconds the speech signal lasts.

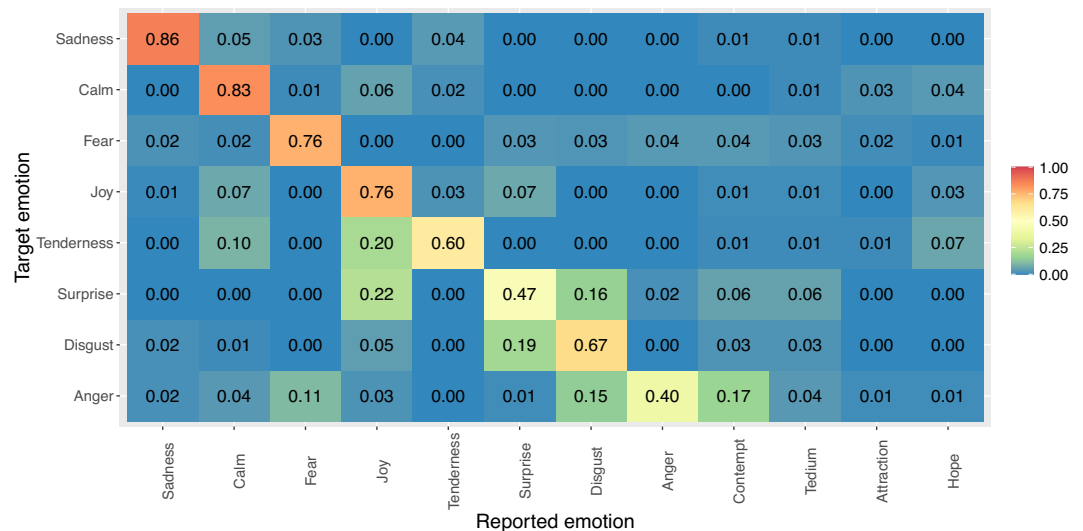


Fig. 2 Self-reported emotion labeling distribution (0.00-1.00) comparing the target discrete emotions to the reported ones. It includes 8 target emotions versus 12 reported ones.

Besides data collection, additional informative documents regarding, for instance, sensor placement during the experiment or the self-assessment instructions, are included in each folder of the data repository.

Technical Validation

Emotional elicitation and labeling is a complex task, and sometimes the expected (or targeted) emotions are not the ones the volunteers experienced (or reported). The agreement between the target class and the self-reported discrete emotion annotations by the volunteers in this experiment is shown in the matrix in Fig. 2, where it is observed as the ratio of times a targeted emotion is identified and felt as such by the volunteers. Thus, a value of 1.00 means a perfect agreement between the targeted emotion and the emotion felt, and 0.00 means no agreement. As introduced before, only 8 of the 12 basic emotions initially selected were included in WEMAC (see the Stimuli Section), although the 12 emotions were considered for the discrete emotion labeling (see the Measures Section). It means that a volunteer could report an emotion, which is outside the target emotion group in WEMAC. This is also the reason because Fig. 2 shows a rectangular matrix. Analyzing this figure it can be found that the non-included emotions (*attraction*, *contempt*, *hope* and *tedium*) are very scarcely selected with the exception of the 17% of times a stimulus expected to represent *anger* is taken as *contempt*. It is also observed that *sadness*, *calm*, *joy* and *fear* are the emotions best identified, being the agreement in the fear emotion especially relevant for the use case. *Tenderness* and *disgust* are also quite well portrayed by the stimuli while *anger* is often taken as *disgust* or *contempt*, and *surprise* as *joy* or *disgust*.

To statistically confirm the associations observed between the target and self-reported discrete emotions, a Pearson's chi-squared test is used as implemented in the *chisq.test()* function in the R software. In this test, the null hypothesis (H_0) states that the variables (i.e., the target and reported emotions) are independent, meaning there is no relationship between categorical variables. As a result of this test, a p-value equaling 0.0005 is obtained, resulting in the dependency between categories is confirmed, with a 0.01 significance level. Then, we perform a study regarding the relationships between the different target and reported individual emotions. To this end, a posthoc analysis with Bonferroni correction for multiple comparisons is performed for each combination of target and reported emotion. This analysis is performed by means of the Pearson standardized residuals (Z-factor) implemented in the *stdres* component of the result of the previous *chisq.test()* function in the R software. The Z-factors appear in Table 3, where a strong dependence occurs when the absolute value of the Z-factor is higher than the reference Z-factor (3.34).

In such a case, a positive Z-factor means a strong direct dependence and a negative one means a strong inverse dependence.

The following conclusions are extracted from analyzing Table 3: (1) all the target emotions show a strong positive dependence with their corresponding reported, so the videos are especially eliciting the desired emotion in the users; (2) all videos originally classified as *non-fear* have a significant negative dependence with such emotion, which implies *fear* is clearly recognizable; (3) it is statistically confirmed that the clips classified as *disgust*, *tenderness* and *anger* show positive dependence with other emotions in addition to their own. Focusing on this last conclusion, it is remarkable the dependence between *disgust* and *surprise* since they are far from each other in the VAD space, meaning they should be easy to differentiate. Note that the VAD emotional space is also known as PAD in the literature. For the videos classified as *anger*, positive correlations are observed with the emotions of *disgust* and *contempt*, which fits with the behavior already obtained in previous studies^{29,32} when it comes to clips that show gender-based violence scenes.

The dimensional VAD annotations are also examined employing the Intra-class Correlation Coefficient (ICC) based on a single-rating, absolute-agreement, two-way mixed effects model, as implemented in the *icc()*

Targeted emotion	Reported emotion											
	Surprise	Anger	Attraction	Calm	Contempt	Disgust	Fear	Hope	Joy	Sadness	Tedium	Tenderness
Surprise	12.95	-1.23	-0.77	-2.25	0.59	2.05	-5.43	-0.92	2.44	-2.06	1.49	-1.63
Anger	-2.38	18.06	0.22	-2.30	8.08	3.40	-6.87	-0.33	-3.60	-2.79	1.00	-2.98
Calm	-2.50	-2.67	1.81	26.71	-2.23	-3.13	-7.69	1.92	-1.79	-3.00	-1.09	-1.42
Disgust	6.14	-2.67	-1.12	-2.91	-0.72	21.99	-7.90	-1.34	-2.11	-2.22	0.18	-2.38
Fear	-3.21	-3.18	1.30	-8.93	-0.71	-6.82	28.32	-2.39	-12.54	-6.84	1.12	-7.98
Joy	1.02	-3.35	-1.40	-0.81	-1.96	3.93	-9.91	1.70*	26.37	-3.34	-1.11	-1.40
Sadness	-2.50	-2.67	-1.12	-1.47	-1.73	-3.13	-7.25	-1.34	-3.74	30.45	-1.09	-0.47
Tenderness	-2.50	-2.67	-0.14	0.34	-1.73	-3.13	-7.90	4.37	2.76	-3.00	-1.09	26.18

Table 3. Pearson standardized residual values (Z-factor) between the target and reported discrete emotions. Values out of the range $[-3.34, 3.34]$ are in bold, meaning strong dependence.

		Surprise	Anger	Attraction	Calm	Contempt	Disgust	Fear	Hope	Joy	Sadness	Tedium	Tenderness
ICC index	Targeted	0.48	0.22	—	0.84	—	0.06	0.70	—	0.52	0.30	—	0.72
	Reported	0.31	0.51	0.40	0.85	0.45	0.26	0.81	0.67	0.56	0.38	0.66	0.69

Table 4. Results of ICC calculation of dimensional VAD annotations using single-rating, absolute-agreement, two-way mixed-random effects model concerning targeted and reported discrete emotions.

function from the library *irr* in the R software. This study is used to evaluate the inter-rater consistency of the dimensional annotations (targeted or reported ones) when classifying stimuli to the corresponding discrete annotations (targeted or reported ones). Based on the 95% confidence interval of the ICC estimate, ICC index values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 are considered poor, moderate, good, and excellent reliability, respectively⁵⁶.

Table 4 shows the reliability ICC metrics analysing the targeted dimensional annotations regarding the targeted discrete ones (see the Targeted field). Analysing this table it is found poor consistency for 4 out of 8 of the emotions, which are *surprise*, *anger*, *disgust*, and *sadness*. It highlights the case of *disgust* with an even almost zero reliability. However, a moderated reliability (close to good) is obtained for *fear* and *tenderness*, getting good reliability for *calm*. This table also shows the reliability ICC metrics analysing the reported dimensional annotations regarding the reported discrete ones (see the Reported field). Analysing this table, it is found that the reliability obtained is slightly better than in the targeted case with poor consistency for 5 out of 12 of the emotions (*surprise*, *attraction*, *contempt*, *disgust*, and *sadness*), a moderated reliability for *anger*, *hope*, *joy*, *tedium*, and *tenderness*, and good reliability for *calm* and *fear*. From this study, we conclude that (1) the dimensional labeling procedure is less robust than the discrete one (it could be due to the difficulty in understanding and applying the VAD metrics); (2) the dimensional labeling for the reported emotions is slightly robust than the targeted one; and (3) fear is one of the most robust labeled dimensional emotions.

Usage Notes

Emotion Recognition. The most straightforward use of this database is the classification of the emotions felt by a woman, using machine and deep learning algorithms with unimodal and multimodal approaches. First, the physiological signals can be used together or separately to analyze their relationship with the annotated discrete or dimensional emotions. The same can be done with the audio signal because even though it can be considered as part of the annotation procedure, it is not acknowledged as such, because its purpose is to have captured the last traces of the emotion elicited in the person. Moreover, the physiological signals were recorded during the entire experiment, so that synchronization can be made with the physiological and audio signals, leading to a multi-modal or fusion scheme. On this basis, a series of experiments carried out in mono- and multi-modal emotion recognition can be found in “Supplementary Material”.

The libraries recommended for further processing of the WEMAC dataset are the ones we found most useful for data cleaning and filtering for physiological and speech signals. On the one hand, Matlab[®] was employed for the physiological data processing using the TEAP toolbox (<https://github.com/Gijom/TEAP>). On the other hand, for the speech signals, the librosa library (<https://librosa.org/doc/latest/index.html>) facilitates the loading of the speech signals and its processing, and for the extraction of the features the openSMILE library (<https://audeering.github.io/opensmile-python/>), DeepSpectrum module (<https://github.com/DeepSpectrum/DeepSpectrum>) and the VGGish module (<https://github.com/tensorflow/models/tree/master/research/audio-sets/vggish/>) are used.

Accessing data - End User License Agreement (EULA). The use of the WEMAC dataset is licensed under a Creative Commons Attribution 4.0 International License (CC-BY-4.0). The data is encrypted and hosted in the UC3M4Safety Repository in the ‘*Consortio Madroño*’ online platform at <https://edatos.consortiomadrono.es/dataverse/empatia>, which includes one folder per dataset described in Table 1. Instructions to decrypt the data will be provided after fulfilling the EULA form located at <https://www.uc3m.es/institute-gender-studies/DATASETS>, which should be signed and emailed to the UC3M4Safety Team (uc3m4safety@uc3m.es).

Code availability

On the one hand, for the physiological signals, as they are raw data, the prospective researchers can work on them with any signal processing toolkit available in software tools (Matlab(R), R, etc.). On the other hand, the code used for the speech data processing and cleaning (developed in Python 3.6.5) is publicly available at https://github.com/BINDI-UC3M/wemac_dataset_signal_processing. All required packages are listed in the repository, which is expected to serve as a starting point for further data analyses.

Received: 31 October 2023; Accepted: 14 October 2024;

Published online: 30 October 2024

References

- Nations, U. Declaration on the elimination of violence against women (1993).
- Sardinha, L., Maheu-Giroux, M., Stöckl, H., Meyer, S. R. & García-Moreno, C. Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. *The Lancet* [https://doi.org/10.1016/S0140-6736\(21\)02664-7](https://doi.org/10.1016/S0140-6736(21)02664-7) (2022).
- Research, Trend Analysis Branch, U. N. O. o. D. & (UNODC), C. Killings of women and girls by their intimate partner or other family members - global estimates 2020 - world (2021).
- Segrave, M. & Vitis, L. (eds.) *Gender, Technology and Violence*. Routledge Studies in Crime and Society (1st edn, Routledge, United Kingdom, 2017).
- Miranda, J. A. *et al.* Bindi: Affective internet of things to combat gender-based violence. *IEEE Internet of Things Journal* (2022).
- Kreibig, S. D. Autonomic nervous system activity in emotion: A review. *Biological psychology* **84**, 394–421 (2010).
- Schmidt, P., Reiss, A., Dürichen, R. & Laerhoven, K. V. Wearable-based affect recognition—a review. *Sensors* **19**, <https://doi.org/10.3390/s19194079> (2019).
- Koolagudi, S. G. & Rao, K. S. Emotion recognition from speech: a review. *International journal of speech technology* **15**, 99–117 (2012).
- Poria, S., Cambria, E., Bajpai, R. & Hussain, A. A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion* **37**, 98–125, <https://doi.org/10.1016/j.inffus.2017.02.003> (2017).
- Zhang, J., Yin, Z., Chen, P. & Nichele, S. Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion* **59**, 103–126 (2020).
- Cimtay, Y., Ekmekcioglu, E. & Caglar-Ozhan, S. Cross-subject multimodal emotion recognition based on hybrid fusion. *IEEE Access* **8**, 168865–168878 (2020).
- Huang, Y., Yang, J., Liu, S. & Pan, J. Combining facial expressions and electroencephalography to enhance emotion recognition. *Future Internet* **11**, 105 (2019).
- Muaremi, A., Bexheti, A., Gravenhorst, F., Arnrich, B. & Tröster, G. Monitoring the impact of stress on the sleep patterns of pilgrims using wearable sensors. In *IEEE-EMBS international conference on biomedical and health informatics (BHI)*, 185–188 (IEEE, 2014).
- Kanjo, E., Younis, E. M. & Sherkat, N. Towards unravelling the relationship between on-body, environmental and emotion data using sensor information fusion approach. *Information Fusion* **40**, 18–31 (2018).
- Koelstra, S. *et al.* Deap: A database for emotion analysis using physiological signals. *IEEE Transactions on Affective Computing* **3**, 18–31 (2012).
- Soleymani, M., Lichtenauer, J., Pun, T. & Pantic, M. A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing* **3**, 42–55 (2012).
- Schmidt, P., Reiss, A., Duerichen, R., Marberger, C. & Van Laerhoven, K. Introducing wesad, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction, ICMi' 18*, 400–408, <https://doi.org/10.1145/3242969.3242985> (Association for Computing Machinery, New York, NY, USA, 2018).
- Correa, J. A. M., Abadi, M. K., Sebe, N. & Patras, I. Amigos: A dataset for affect, personality and mood research on individuals and groups. *IEEE Transactions on Affective Computing* (2018).
- Baird, A. *et al.* An evaluation of speech-based recognition of emotional and physiological markers of stress. *Frontiers in Computer Science* **3**, <https://doi.org/10.3389/fcomp.2021.750284> (2021).
- Baird, A., Amiriparian, S., Berschneider, M., Schmitt, M. & Schuller, B. Predicting biological signals from speech: Introducing a novel multimodal dataset and results. In *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, 1–5, <https://doi.org/10.1109/MMSP.2019.8901758> (2019).
- Robinson, D. L. Brain function, emotional experience and personality. *Netherlands Journal of Psychology* **64**, 152–168, <https://doi.org/10.1007/BF03076418> (2008).
- Blanco Ruiz, M. *et al.* UC3M4Safety Database description. <https://edatos.consorciomadrone.es/dataverse/empatia> (2021).
- Blanco Ruiz, M. *et al.* UC3M4Safety Database - List of Audiovisual Stimuli (Video), <https://doi.org/10.21950/LUO1IZ> (2021).
- Blanco Ruiz, M. *et al.* UC3M4Safety Database - List of Audiovisual Stimuli, <https://doi.org/10.21950/CXAAHR> (2021).
- Miranda Calero, J. A. *et al.* UC3M4Safety Database - WEMAC: Biopsychosocial questionnaire and informed consent, <https://doi.org/10.21950/U5DXJR> (2022).
- Miranda Calero, J. A. *et al.* UC3M4Safety Database - WEMAC: Physiological signals, <https://doi.org/10.21950/FNUHKE> (2022).
- Rituerto González, E. *et al.* UC3M4Safety Database - WEMAC: Audio features, <https://doi.org/10.21950/XKHCCW> (2022).
- Miranda Calero, J. A. *et al.* UC3M4Safety Database - WEMAC: Emotional labelling, <https://doi.org/10.21950/RUUCLV> (2022).
- Sainz-de Baranda Andujar, C., Gutiérrez-Martín, L., Miranda-Calero, J. Á., Blanco-Ruiz, M. & López-Ongil, C. Gender biases in the training methods of affective computing: Redesign and validation of the self-assessment manikin in measuring emotions via audiovisual clips. *Frontiers in Psychology* **13**, 955530 (2022).
- Stanney, K., Fidopiastis, C. & Foster, L. Virtual reality is sexist: But it does not have to be. *Frontiers in Robotics and AI*, **7**, <https://doi.org/10.3389/frobt.2020.00004> (2020).
- Saredakis, D. *et al.* Factors associated with virtual reality sickness in head-mounted displays: A systematic review and meta-analysis. *Frontiers in Human Neuroscience* **14**, <https://doi.org/10.3389/fnhum.2020.00096> (2020).
- Blanco-Ruiz, M., Sainz-de Baranda, C., Gutiérrez-Martín, L., Romero-Perales, E. & López-Ongil, C. Emotion elicitation under audiovisual stimuli reception: Should artificial intelligence consider the gender perspective? *International Journal of Environmental Research and Public Health* **17**, <https://doi.org/10.3390/ijerph17228534> (2020).
- Fontaine, J. R., Scherer, K. R., Roesch, E. B. & Ellsworth, P. C. The world of emotions is not two-dimensional. *Psychological Science* **18**, 1050–1057, <https://doi.org/10.1111/j.1467-9280.2007.02024.x> (2007).
- Russell, J. & Mehrabian, A. Evidence for a three-factor theory of emotions. *Journal of Research in Personality* **11**, 273–294, [https://doi.org/10.1016/0092-6566\(77\)90037-X](https://doi.org/10.1016/0092-6566(77)90037-X) (1977).
- Liu, H., Hartmann, Y. & Schultz, T. Csl-share: A multimodal wearable sensor-based human activity dataset (2021).
- Liu, H., Hartmann, Y. & Schultz, T. A practical wearable sensor-based human activity recognition research pipeline. In *HEALTHINF*, 847–856 (2022).

37. Carvalho, M. & Brás, S. Heartbeat selection based on outlier removal. In *Iberian Conference on Pattern Recognition and Image Analysis*, 218–229 (Springer, 2022).
38. Harjani, M., Grover, M., Sharma, N. & Kaushik, I. Analysis of various machine learning algorithm for cardiac pulse prediction. In *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 244–249 (IEEE, 2019).
39. Gutiérrez Martín, L. Entorno de entrenamiento para detección de emociones en víctimas de Violencia de Género mediante realidad virtual. In *Bachelor Thesis* (University Carlos III de Madrid, 2019).
40. Miranda, J. A., Canabal, M. F., Portela García, M. & Lopez-Ongil, C. Embedded emotion recognition: Autonomous multimodal affective internet of things. In *Proceedings of the cyber-physical systems workshop*, vol. 2208, 22–29 (2018).
41. Miranda, J. A., Canabal, M. F., Gutiérrez-Martín, L., Lanza-Gutiérrez, J. M. & López-Ongil, C. A design space exploration for heart rate variability in a wearable smart device. In *2020 XXXV Conference on Design of Circuits and Integrated Systems (DCIS)*, 1–6, <https://doi.org/10.1109/DCIS51330.2020.9268628> (2020).
42. Canabal, M. F., Miranda, J. A., Lanza-Gutiérrez, J. M., Pérez Garcilópez, A. I. & López-Ongil, C. Electrodermal activity smart sensor integration in a wearable affective computing system. In *2020 XXXV Conference on Design of Circuits and Integrated Systems (DCIS)*, 1–6, <https://doi.org/10.1109/DCIS51330.2020.9268662> (2020).
43. Miranda Calero, J. A., Páez-Montoro, A., López-Ongil, C. & Paton, S. Self-adjustable galvanic skin response sensor for physiological monitoring. *IEEE Sensors Journal* **23**, 3005–3019, <https://doi.org/10.1109/JSEN.2022.3233439> (2023).
44. Rottenberg, J., Ray, R. & Gross, J. Emotion elicitation using films in: Coan ja, allen jib, editors. the handbook of emotion elicitation and assessment (2007).
45. McFee, B. *et al.* librosa, audio and music signal analysis in python. In *Proceedings of the 14th python in science conference*, 18–24, <https://doi.org/10.25080/Majora-7b98e3ed-003> (2015).
46. Stappen, L. *et al.* The muse 2021 multimodal sentiment analysis challenge: Sentiment, emotion, physiological-emotion, and stress (2021). 2104.07123.
47. McFee, B. *et al.* librosa/librosa: 0.9.1, <https://doi.org/10.5281/zenodo.6097378> (2022).
48. Eyben, F. *et al.* The geneva minimalist acoustic parameter set (gemaps) for voice research and affective computing. *IEEE Transactions on Affective Computing* **7**, 1–1, <https://doi.org/10.1109/TAFFC.2015.2457417> (2015).
49. Eyben, F., Wöllmer, M. & Schuller, B. opensmile – the munich versatile and fast open-source audio feature extractor. In *Proceedings of the 18th ACM international conference on Multimedia*, 1459–1462, <https://doi.org/10.1145/1873951.1874246> (2010).
50. Schuller, B. *et al.* The interspeech 2016 computational paralinguistics challenge: Deception, sincerity & native language. 2001–2005, <https://doi.org/10.21437/Interspeech.2016-129> (2016).
51. Eyben, F. Real-time Speech and Music Classification by Large Audio Feature Space Extraction. PhD Thesis, Technische Universität München, München (2015).
52. Amiriparian, S. *et al.* Snore sound classification using image-based deep spectrum features. In *Interspeech 2017*, 3512–3516 (ISCA, 2017).
53. Russakovsky, O. *et al.* ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)* **115**, 211–252, <https://doi.org/10.1007/s11263-015-0816-y> (2015).
54. Gemmeke, J. F. *et al.* Audio set: An ontology and human-labeled dataset for audio events. In *Proc. IEEE ICASSP 2017* (New Orleans, LA, 2017).
55. Ravanelli, M. *et al.* Multi-task self-supervised learning for robust speech recognition, <https://doi.org/10.48550/ARXIV.2001.09239> (2020).
56. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine* **15**, 155–163, <https://doi.org/10.1016/j.jcm.2016.02.012> (2016).

Acknowledgements

This work has been supported by the Dept. of Research and Innovation of Madrid Regional Authority, in the EMPATIA-CM research project (reference Y2018/TCS-5046), SAPIENTAE4Bindi Grant PDC2021-121071-I00 funded by MICIU/AEI/10.13039/501100011033 and by the European Union “NextGenerationEU/PRTR”, grant PID2021-125780NB-I00 Grant PID2021-125780NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by “ERDF/EU”, and the Spanish Ministry of Science, Innovation and Universities with the FPU grant FPU19/00448, and by the National Research Project Bindi-Tattoo PID2022-142710OB-I00 (Agencia Estatal de Investigación, AEI, Spain). The authors thank all the members of UC3M4Safety for their contribution and support, especially Rosa San Segundo, Clara Sainz de Baranda Andujar, Marian Blanco Ruiz, Eva Martinez Rubio, and Manuel Felipe Canabal for their contribution to the data-gathering, experimental protocol design and volunteer screening.

Author contributions

Conceptualization and Design of the Database: All authors. Database Capture Protocol: All authors. Participants Assistance: C.L.O., E.R.G., E.R.P., J.M.C. and L.G.M. Data Curation: E.R.G., E.R.P., J.M.C. and L.G.M. Technical Validation: E.R.P. and L.G.M. Supplementary Material Experimentation: E.R.G. and J.M.C. Original Draft Writing: E.R.G., E.R.P., J.L.G., J.M.C. and L.G.M. Supervision: C.P.M. and C.L.O. Review & Editing: J.L.G., C.P.M. and C.L.O.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-024-04002-8>.

Correspondence and requests for materials should be addressed to J.A.M. or E.R.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024