# An intelligent and generic approach for detecting human emotions

Mano, Leandro Y.; Faical, Bruno S.; Goncalves, Vinicius P.; Pessin, Gustavo; Gomes, Pedro H.; de Carvalho, Andre C. P. L. F.; Ueyama, Jo

*Published in:*
Applied Soft Computing

*DOI:*
10.1007/s00500-019-04411-7

Soft Computing (2020) 24:8467–8479
https://doi.org/10.1007/s00500-019-04411-7

METHODOLOGIES AND APPLICATION

# An intelligent and generic approach for detecting human emotions: a case study with facial expressions

**Leandro Y. Mano**[1,2] · **Bruno S. Faiçal**[1] · **Vinícius P. Gonçalves**[3] · **Gustavo Pessin**[4] · **Pedro H. Gomes**[5] · **André C. P. L. F. de Carvalho**[1] · **Jó Ueyama**[1]

Published online: 8 October 2019


**Abstract**

Several studies in the field of human–computer interaction have focused on the importance of emotional factors related to the interaction of humans with computer systems. According to the knowledge of the users' emotions, intelligent software can be developed for interacting and even influencing users. However, such a scenario is still a challenge in the field of human–computer interaction. This article endeavors to enhance intelligence in such types of systems by adopting an ensemble-based model that is able to identify and classify emotions. We developed a system (music player) that can be used as a mechanism to interact and/or persuade someone to "change" his/her current emotional state. In order to do this, we also designed a generic model that accepts any kind of interaction or persuasion mechanism (e.g., preferred YouTube channel videos, games, etc.) to be deployed at runtime based on the needs of each user. We showed that the approach based on a genetic algorithm for the weight assignment of the ensemble achieved an accuracy average of 80%. Moreover, the results showed a 60% increase in the level of user's satisfaction regarding the interaction with users' emotions.

**Keywords** Human–computer interaction · Emotion classification · Ensemble of classifiers · Genetic algorithm · Generic approach

## 1 Introduction

Emotion is an important aspect of human interaction with the environment. It plays an important role in interpersonal relationships (Zhou et al. 2011) and it has been extensively studied in the field of psychology (Lichtenstein et al. 2008). The number of studies on users' emotions in the area of human–computer interaction (HCI) has significantly increased over the past years (Klein et al. 2002; Bailenson et al. 2008; Zhou et al. 2011; Peter and Urban 2012; LiKamWa et al. 2013). They have explored the way computing devices can recognize, model and respond to human emotions (among other factors) and how such emotions can be expressed through computer interfaces or interactions (Gonçalves et al. 2017a).

Nonetheless, emotions are complex and hard to identify or evaluate and are linked to several factors, such as motor expressions (Scherer 2005; Mahlke and Minge 2008). Most research projects are concentrated on the use of machine learning algorithms for emotional recognition and treatment of emotions (Libralon and Romero 2014; Lee et al. 2015; Kalsi and Rai 2017), physiological signals (Zhou et al. 2011; Peter and Urban 2012) or behavioral aspects (LiKamWa et al. 2013) which has given rise two basic questions: what are emotions? and what should be studied for the determination of their meaning? Studies have addressed the importance of emotional aspects in the interaction with computational systems (Gonçalves et al. 2017a). The main objective of promoting this affective interaction is to contribute to which

Communicated by V. Loia.

✉ Leandro Y. Mano
leandroyukiomano@usp.br

1 Institute of Mathematical and Computer Sciences, University of Sao Paulo, Sao Carlos, Sao Paulo, Brazil

2 Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Groningen, The Netherlands

3 Department of Electrical Engineering, University of Brasília, Brasília, DF, Brazil

4 Robotics Lab, Vale Technology Institute, Ouro Prêto, MG, Brazil

5 Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA

systems are capable of recognizing the user's emotions, interpreting them and reacting intelligently and responsibly and can best meet the interaction requirements of individuals.

This paper addresses this question with a system that (i) intelligently detects user emotions, and (ii) suggests, through a music player, changes in their real-time behavior. In our approach, the emotion detection is implemented by machine learning techniques that identify and classify users' emotions, based on their motor expressions, particularly in the face. FaceTracker (Saragih et al. 2011) tracks and maps the face feature points, and an approach based on Ensemble of Classifiers (EC) (Mano 2018) classifies and identifies their emotions. Thus, an analysis of the user's face can be conducted and then the HCI can be applied, for example, by suggesting a pleasant task that could be carried out when he/she is under stress (Gonçalves et al. 2017b).

The model addresses a type of interactive method, for example the users' favorite music track and tuning into their favorite YouTube channels. This is proposed in order to provide the possibility that the system adapts according to the use and the preferences of the user, being able to provide several types of platforms of emotional interaction. An analysis of the algorithm was conducted toward validating the model and involved a music player during the user's interaction. The model is more promising than other approaches available in the literature, and the interaction with the music player considerably increased the user's satisfaction.

In this sense, we emphasize the importance of emotional interaction to increase coherence, consistency and credibility of reactions and computational responses during human interaction, through a human–computer interface.

The main contributions of this research involve:

– a discussion on the state-of-the-art studies considering the identification of emotions;
– an approach through which the user's face is used for interaction with an intelligent emotion recognition system;
– an evaluation of an approach based on EC for the classification of user's emotion, together with a Genetic Algorithm (GA) for the determination of weights assigned to each algorithm that composes the Ensemble technique; and
– an approach for interaction of the emotional state.

The article is structured as follows: Sect. 2 addresses some underlying fundamental concepts in the frame of the use of the face for emotion identification; Sect. 3 is devoted to a literature review on relevant topics in emotion recognition and classification systems; Sect. 4 describes face mapping; Sect. 5 focuses on the emotion classification module and our generic software for user interaction; Sect. 6 reports on the assessment of the emotion classification system and the results of our generic approach for emotion interaction; Sect. 7 provides the conclusions and suggests some future work. Table 1 shows the nomenclature/abbreviations used.

**Table 1** List of abbreviations and acronyms

| Abbreviations | Full description |
|---|---|
| CK+ | Extended Cohn–Kanade |
| EC | Ensemble of classifiers |
| GA | Genetic algorithm |
| HCI | Human–computer interaction |
| kNN | *k*-Nearest neighbor |
| ML | Machine learning |
| RaFD | Radboud faces |
| SVM | Support vector machine |

## 2 Previous knowledge of emotions—a psychological view

The emotion recognition task is based on psychological and cognitive science (Russell 1980), which involves two basic views on the representation of emotions, namely (i) continuous and (ii) categorical. The continuous view describes emotions as points in multidimensional space represented on continuous scales or bases for dimensional spaces. Therefore, an emotional state can be characterized as a small number of latent dimensions rather than a small number of emotional categories. According to this standpoint, affective states are not discrete and independent, but systematically related to each other. The most common dimensional approach is the Circumplex Model of Affect, devised by Russell (1980) (see Fig. 1). In the frame of this model, all emotions lie in a two-dimensional space continuum, where dimensions are valence (*X*-axis—degree at which emotions are positive or negative) and excitement (*Y*-axis—energy or excitation level caused by the emotion).

In the categorical representation, different emotions are mapped into distinct categories. Motor expressions communicate the behavioral tendencies of an individual (Scherer 2005; Mahlke and Minge 2008), of which the most well-known example is the set of six universal basic emotions, namely: happiness, disgust, fear, anger, surprise and sadness (Ekman et al. 1987; Russell 1994), and the facial expressions related to them. Such an emotional feature involves changes in facial expressions that register the users' emotional experience (Gonçalves et al. 2017c). The face undergoes changes that reflect the extent to which excitement has arisen; emotional responses, for example, involve a look of hatred, a compression of the lips or even a smile. All other emotional categories are built from combinations of the basic emotions.

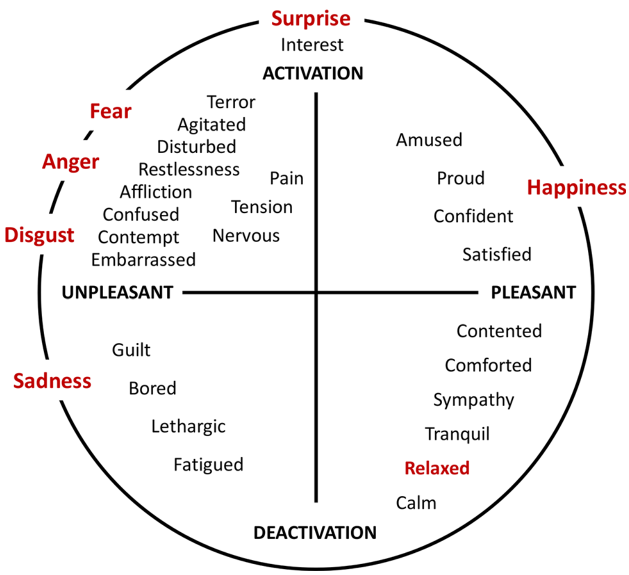


**Fig. 1** Russel's circumplex model (Russell 1980)


**Fig. 2** Basic emotions proposed by Ekman et al. (1987) and Russell (1994). The person in the photos is the user submitted to one experiment

The approach is supported by the cross-cultural studies conducted by Ekman et al. (1987), which claim humans perceive certain basic emotions conveyed by facial expressions in the same way, regardless of culture and ethnicity.

The main advantage of representation in the form of categories is that individuals can use this categorical system to describe emotions observed in daily life. Figure 2 shows the emotions, whose analyses were based on the model proposed by Ekman et al. (1987) and Russell (1994). Neutral state is also taken into account, since it is used as a reference point for the detection of emotional states.

Both the continuous and the categorical representation of emotions have been used in several computational applications and achieved good performance (Scherer 2005; Libralon and Romero 2014; Mano et al. 2016a). The wide spectrum of applications and the increasing amount of computing processing capacity have motivated researchers to identify users' emotions in computing systems and use the information as a basis for decision-making (Lee et al. 2015; Mano et al. 2016; Kalsi and Rai 2017).

## 3 Related work

Research in the area of pattern recognition has focused on techniques for extraction of decisions. One of the application domains is the classification of emotional responses. In classification tasks, Machine Learning (ML) models, through models from past data, predict the class label for data whose classes are unknown (Zhou et al. 2011; Peter and Urban 2012; LiKamWa et al. 2013; Libralon and Romero 2014; Lee et al. 2015; Kalsi and Rai 2017).

An example is the use of physiological sensors for the detection and classification of emotion, through measurements of the user's physiological reactions. Zhou et al. (2011) compared the predictive performance of three classification models for classifying emotional states (excited, amused, contented, neutral, sad, fearful and disgusted) based on physiological reactions. An experiment provoked emotional responses to standardized photos in emotional situations, so that the physiological responses of the participants could be measured. Three classification algorithms, namely: Decision Rule Induction, *k*-Nearest Neighbor (kNN) and Decomposition Tree Induction, were applied to the data set for the construction of predictive models able to predict an emotional situation.

Similarly, Peter and Urban (2012) adopted a general approach to emotion recognition (happiness, euphoria, surprise, stress, anger, fear and boredom) that uses those sensors. The authors designed the so-called emotions recognition system, on the basis of pattern recognition methods (kNN, regression tree, Bayesian networks, support vector machines (SVM), fuzzy logic, artificial neural networks), which collects data such as electrodermal activity, atmospheric air temperature and heart rate variability from some types of sensors.

LiKamWa et al. (2013) designed an approach that detects and classifies emotions through observation of trends in user's behavior. A model recognizes humor (positive and negative) analyzing communication history and application usage patterns expressed in six types of information, namely: posts, e-mail, log connections, applications used, web browsing, communication and location. It differs from other mood detection applications because it does not rely on physical sensors embedded in devices.

A study conducted by Libralon and Romero (2014) sets out a model based on classification algorithms—linear regression and artificial neural networks of type multilayer perceptron—for real-time analysis of emotions (happiness, fear, neutral, anger, surprise and sadness) based on motor expressions. Libralon and Romero (2014) investigated the recognition of emotions obtained by an ensemble of features of facial expressions. Features, such as shapes of the facial components (e.g., eyes, mouth, nose, chin and eyebrows) and location of salient facial parts (e.g., corners of the eye and

**Table 2** Summary of the papers identified in the literature

| Related work references | Emotional component | Number of emotions | Classification technique | User interaction |
|---|---|---|---|---|
| Zhou et al. (2011) | Physiological | 7 | Machine learning | X |
| Peter and Urban (2012) | Physiological | 7 | Machine learning | X |
| LiKamWa et al. (2013) | Behavioral trend | 2 | History and usage | X |
| Libralon and Romero (2014) | Expressions | 6 | Machine learning | X |
| Lee et al. (2015) | Expressions | 5 | Distance from Hausdorff | X |
| Kalsi and Rai (2017) | Expressions | 2 | Machine learning | X |
| Our proposal | Expressions | 7 | Machine learning | √ |

mouth) are used for distinguishing six different facial elements that reveal facial changes. The authors showed that the use of predefined areas of the face, together with angles and distances, is a valid way of building emotional classification models.

Lee et al. (2015) designed an approach that automatically segments an image and recognizes emotions (happy, neutral, anger, surprise and sadness) interpreting facial expressions to extract characteristic points and the variation of the Bezier curve of the still image. The proposal suggests three steps: (1) detection of facial regions with a map of characteristics, (2) Bezier curve in the eyes and mouth, and (3) classification of the emotion of the characteristic with Hausdorff distance.

Finally, Kalsi and Rai (2017) developed a classification system that recognizes emotions of the expressions of user's face images. The proposal addresses the preprocessing of a facial image followed by the extraction of characteristics used to construct classifiers for two elementary (happy and sad) emotional states. This work uses the binary method of location of the approximation image to extract features along with the SVM algorithm to classify the emotion.

Table 2 shows an overview of studies found in the literature and detailed above. The research was analyzed considering the following characteristics: (i) emotional component for the classification of emotion; (ii) number of emotions classified; (iii) technique used for the classification of emotion; and (iv) user tests for interaction.

Some recent optimization approaches have addressed real problems in the search for efficiency of the applied algorithms. Shubham and Kusum (2019a) introduced the gray wolf optimizer algorithm for the field of swarms intelligence to solve continuous problems of optimization. The results showed efficiency and reliability of the algorithm in solving continuous optimization problems and real-life optimization problems. Shubham and Kusum (2019b, c) improved the Sine Cosine algorithm based on the characteristics of the sine and cosine trigonometric functions. The algorithm increases the exploitation ability of solutions and reduces the diversity present in search equations. Its main characteristic is the hybridization of crossover exploration skills with the best personal state of individual solutions and the integration of global search mechanisms and self-learning. The results showed that the algorithm can efficiently solve real-life optimization problems.

An increasing number of studies on emotional classification have identified emotions through sensors and/or information from computing devices and demonstrated that emotional classification promotes more accurate decisions in interactions between users and computing devices. However, despite the frequent use of machine learning techniques for the classification of emotions, the literature lacks studies that use the EC concept and optimization for emotional classification. An EC is evaluated and can fill this gap in the literature and mitigate the resulting errors when using only a single technique to classify the emotions. The model was compared with traditional approaches used by the studies discussed above to classify the emotions, namely: decision tree (D. Tree), naive Bayes (N. Bayes), fuzzy logic, kNN and SVM. Different from the research presented, our work provides, through a music player, an interaction, i.e., for users to interact with computing devices according to their emotions.

## 4 Extraction approach for face identification and expression analysis

According to Tian et al. (2011), the analysis of automatic facial expressions can be divided into three key stages, namely: (i) face detection, (ii) facial feature extraction and (iii) emotion classification. Face detection automatically finds the facial region on the input images. Detecting a face in a complex scene is a non-trivial problem, since head movements, occlusion, changes in the illumination settings and presence of hair or glasses can occur to the system (Tian et al. 2011).

The extraction approach for face identification and expression analysis adopted here is based on geometric features. Methods based on geometric features are used in facial modeling (motor expressions) with a view to adopting an approach which resembles the way that human beings inter-

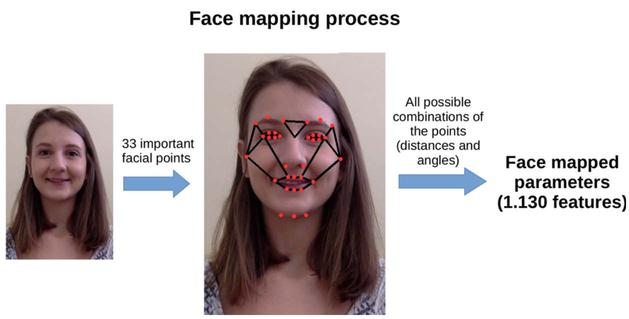


**Fig. 3** Face mapping process by FaceTracker; photos used of the user submitted to the experiment

pret the different parts of the face. Therefore, both shape and locations of facial components (including mouth, eyes, eyebrows and nose) can be modeled through the use of feature points and geometric elements, such as angles, distances or areas that represent the geometry of the face. The classification of facial expressions is the last stage, and ML-based tackle the problem of classification.

FaceTracker (Saragih et al. 2011) is a computer vision system that obtains information about facial features. It uses an optimization strategy that employs a linear approximation by adjusting the reference points in consistent locations to record a model designed to work within fixed parameters. Moreover, the system is based on a face model as a benchmark of 66 feature points. The algorithm aims to align the elements of the face being analyzed with the feature points of the reference model.

We used the approach of Mano et al. (2015) and Mano (2018) for the FaceTracker algorithm, since it maps a subset of the 66 points initially obtained and uses only 33 feature points. The purpose is to reduce computational costs by eliminating possible redundancy. Figure 3 shows an example of a face mapped by FaceTracker.

This process of facial mapping is supported by theories developed by psychologists (Ekman et al. 1987; Mahlke and Minge 2008; Russell 1994; Scherer 2005). As such, we considered (i) eight points that map the mouth, (ii) six for each individual eye, (iii) three for each eyebrow; (iv) three for the chin, (v) two for the nostrils, and (vi) two for the delineation of the lateral extremities of the face near the eyes.

As in Mano et al. (2015) and Mano (2018), we also considered the distances between two distinct points and the angles made by the line connecting them with the horizontal axis, all of them obtained in all possible combinations of the points. As a result, a representation of dimensionality $D_1 = 2 \bullet 33 + 8 + 2 \bullet 528 = 1.130$ is created. Although our model is based on a limited number of facial points (i.e., only 33 facial points), provides a wide number of features (i.e., 1.130, whereas the previous FaceTracker produces or generates 66 features), such features are used for training purposes of our model for emotion classification.

## 5 Our emotion classification and generic approach to emotional interaction

Due to the fact that scientific literature is seriously concerned with searching for alternative methods that can be used to identify and classify user's emotions, we have set out a model based on EC to identify and classify emotions by analyzing the user's face. The model uses a face recognizer (FaceTracker) and identifies and categorizes emotions through a combination of response values of classification algorithms, so that computing systems can interact with the user's emotional state more assertively. Below is a discussion of the EC approach.

### 5.1 Ensemble of classifiers approach for emotion classification

Despite their extensive use, the classifiers generated by ML methods rarely achieve 100% accuracy (Mano et al. 2015; Mano 2018), since their performance depends on several parameter settings, the completeness of the data sample used for the training, and the degree of difficulty associated with the particular problem.

The selection of a single classifier involves rejection of a significant amount of potentially useful information. Therefore, the concept of EC has been regarded as a possible solution for the development of high-performance systems in the area of pattern recognition (Schuller et al. 2005; Mano et al. 2015, 2016b).

A factor that ensures good performance of EC is the range of its components (Duda et al. 2012). Classifiers that compose an EC are assumed to be different from each other if they have uncorrelated errors. For an ensemble to achieve acceptable performance, it must be formed from classifiers of similar and reasonable degree of accuracy and avoid coincidence errors so that the errors of a classifier can be corrected by the choices made by all the other components (Schuller et al. 2005; Duda et al. 2012; Mano et al. 2015, 2016b). For this reason, for our proposal classification techniques were used that have been increasingly employed to analyze the user's emotional responses, such as: kNN, SVM, fuzzy Logic, decision or regression tree and Bayesian networks. Figure 4 shows the module structure for the face, classified by the EC approach.

The first processing layer (layer 1 of Fig. 4) receives the results of the facial mapping performed by FaceTracker as input. It consists of individual classifiers of different architectures; however, they produce outputs in a common way, such as ranking of candidate classes. The second processing layer (layer 2 of Fig. 4) consists in a decision-making process that

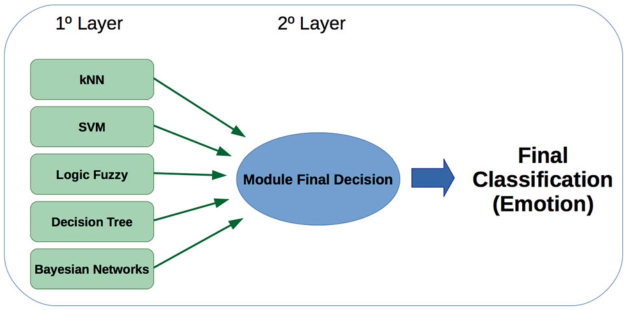


**Fig. 4** Architecture of our approach: ensemble of classifiers for the detection emotions on the user's face

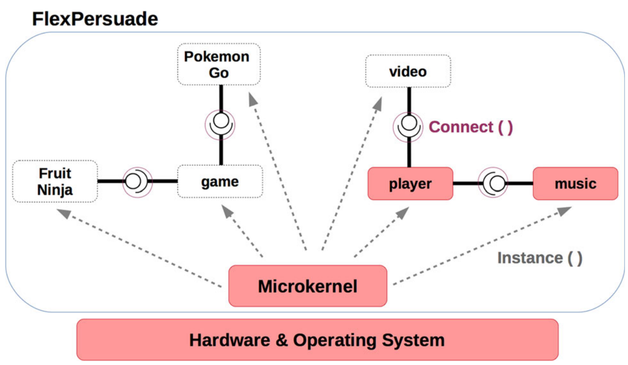


**Fig. 5** Relationship among components on *FlexPersuade*

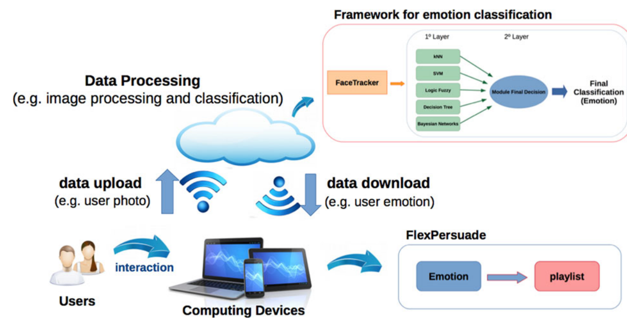


**Fig. 6** Architecture proposed by Mano et al. (2016a, c) for interaction with the computational devices

processes the outputs of the previous layers and forwards a decision of the EC (see Fig. 4). For each algorithm, a weight will be assigned by the genetic algorithm approach for the final decision about the ensemble.

The GA-based approach was adopted to set new values (regarded as optimized values) in response to the weights of each classification algorithm. GA handles data structures to model an evolve behavioral patterns between its elements [the behavior is based on Darwin et al. (1872)] and to find possible solutions to complex problems.

## 5.2 Our generic approach for emotional interaction

A generic approach was adopted for the construction of adaptive applications to computing devices. "Generic," in this case, refers to the possibility for our approach to build several types of emotional interaction platforms, such as music players, preferred YouTube channels and games. As addressed in Sect. 1, emotional interaction can be defined as the actions taken to alter someone's emotional state. In our platform, the minimal kernel (i.e., microkernel) loads and unloads components and connects and disconnects them (see Fig. 5) prior to running the interactive system.

The *FlexPersuade* software was designed according to OpenCom model (Coulson et al. 2008) in which the microkernel entity loads and interconnects the components required for the interaction with the user. In the case of this work, *FlexPersuade* instantiates and connects a music player (on the right of Fig. 5), but it can also be used for other emotional interaction software, e.g., games (on the left of the same illustration).

Figure 6 depicts the structure of the interactions between *FlexPersuade* and the user when a computational device is used. The framework includes a FaceTracker module (Saragih et al. 2011), which extracts the user's facial features, and the ensemble of classifiers algorithm, which classifies an emotion that is based on the features processed in the identification of the user's face (see Sect. 4). Our case study, which involves a music player, *FlexPersuade* loads the playlist that best suits the user's current emotional state (Mano et al. 2016c). For example, if the framework ranks the user's emotional experience as "sad", the player loads a playlist best suited to this feeling, or even tries to "persuade" to lower the level of sadness and change his/her emotional state.

When connecting to the music player, the user should inform on their musical preferences, so that the software download songs that best fit their emotional state. The user fills a digital form to rank the artists he/she usually listens to, according to their emotional state. According to the musical preferences and image of the face, the player plays the songs that must adapt to the user's emotional state. As the images are sorted every 5 s, the emotions are stored in a temporary file and, at the end of each song, the system detects the most frequent emotion during the time interval and the next song to be played, according to the preferences.

# 6 Evaluation of the emotion classification

The experiments reported in this section were performed on two different occasions. First, the user's emotions classification accuracy was evaluated for the model. Images of emotions from different databases, as well as the cross-validation technique were used to fulfill these objectives.

**Table 3** Parameters used in the implementation of the selected algorithms

| Classifiers | Implementation | Parameters |
|---|---|---|
| Naive Bayes | Naive Bayes | debug = false, displayModelInOldFormat = false, useKernelEstimator = false, useSupervisedDiscretization = false |
| Decision tree | J48 | binarySplits = false, collapseTree = true, confidenceFactor = 0.25, debug = false, minNumObj = 2, numFolds = 3, reducedErrorPruning = false, saveInstanceData = false, seed = 1, subtreeRaising = true, unpruned = false, useLaplace = false, useMDLcorrection = true |
| Fuzzy | FuzzyNN | KNN = 10, debug = false, fuzzifier = 3.0, similarity = SimilarityFNN − R first-last − T weka.fuzzy.tnorm.TNormLukasiewicz − C 0.0 |
| kNN | IBk | KNN = 1, crossValidate = false, debug = false, distanceWeighting = No distance weighting, meanSquared = false, nearestNeighbourSearchAlgorithm = LinearNNSearch − A “weka.core.EuclideanDistance − R first-last”, windowSize = 0 |
| SVM | SMO | buildLogisticModels = false, c = 1.0, checksTurnedOff = false, debug = false, epsilon = 1.0E-12, filterType = Normalize training data, kernel = PolyKernek − C 250007 − E 1.0, numFolds = − 1, randomSeed = 1, toleranceParameter = 0.001 |

The second evaluation measured user’s satisfaction when using songs with and without the proposed approach. It consisted in a user performing an activity of interest on a laptop (e.g., using social networks, reading news and writing e-mail) while listening to music indicated by the music player according to the user’s preferences.

### 6.1 Overview

For the training/evaluation of the model pictures of facial expressions with different emotions were used. The databases used were Radboud Faces (RaFD) (Langner et al. 2010), Extended Cohn–Kanade (CK+) (Lucey et al. 2010), IMPA-FACE3D (Mena et al. 2011) and FACES (Ebner et al. 2010), all with open access to the public. RaFD provides examples of facial expressions of 67 American and Moroccan participants (adults and children, male and female). CK+ includes 593 facial expressions of 123 different adults (69% female and 31% male), of whom 81% are Europeans or Americans, 13% are African–Americans and 6% belong to other ethnic groups. IMPA-FACE3D includes pictures of 38 individuals, with samples of the six universal facial expressions and neutral expressions. It consists of 22 men and 16 women, most aged between 20 and 50. Finally, FACES includes images of faces of 171 individuals (58 young, 56 middle-aged and 56 elderly men and women), with the seven facial expressions analyzed in this work.

We selected images of participants facing forward and expressing the following emotions: joy, disgust, fear, anger, surprise and sadness or a neutral state. 67 examples of each emotion were obtained from RaFD, 38 were provided by IMPA-FACE3D and 171 and 169 were obtained from FACES and CK+, respectively, including children, adults and elderly people. We used 3115 images, divided into 445 images for each emotion group for achieving an optimal balance between the training images used for the tests. The combination of the selected images provided a set of heterogeneous tests that included images from different regions and a model to be applied to any computing device regardless of the individual’s ethnicity.

Table 3 shows the parameters used for both individual tests [WEKA Framework (Bouckaert et al. 2013)] and development of the Ensemble model.

We implemented the algorithms and the EC decision module (described in Sect. 5) to assess the accuracy of the classification module. The genetic algorithm was adopted for setting the weights that constitute the EC. As a result, data on the behavior and evaluation of the classification module were collected, and the module was evaluated in a case study in which a music player offers a playlist to stimulate the user according to the emotion identified. The model takes the user’s facial features as input for the automated emotion classification.

### 6.2 Evaluation of the weights used in the ensemble

The optimization problems show an NP-hard level of complexity, which significantly increases their cost to resolve using exact techniques. Therefore, evolutionary computation techniques are adopted for modeling the evolutionary behavior (Darwin and Wallace 1858). According to this theory, individuals belonging to a population with a greater ability

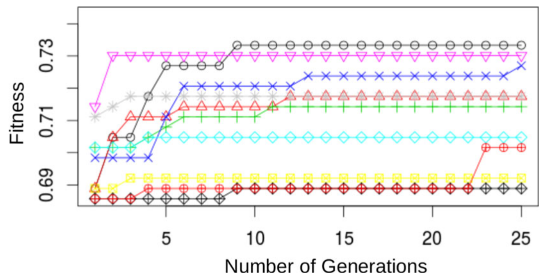


**Fig. 7** Evolution with I12G25 setting—evolution provided by 12 individuals and evolutions of 25 generations

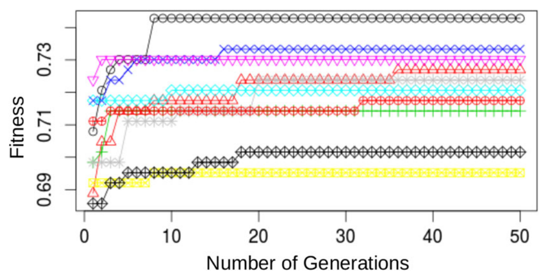


**Fig. 8** Evolution with I25G50 setting—evolution provided by 25 individuals and evolutions of 50 generations

to adapt to a given environment have higher chance of reproducing their characteristics in future generations (Carvalho et al. 2011).

Genetic algorithms (GAs) can be classified as search algorithms modeled in the process of natural evolution, and they can efficiently optimize functions, such as indicating the weights of neurons in an artificial neural network. This approach allows to efficiently identify combinations of weights (solutions) in large and complex search spaces.

For the abovementioned reasons, a GA approach was selected to optimize the weights used in Ensemble. Figures 7, 8 and 9 show the best fitness found during the evolution of the GA in the cross-validation of the calibration group. The GA settings related to the population and generations are written as "**IXGY**", where **X** is the number of individuals, and **Y** is the extent to which generation GA "evolves" the population. Further details about setting GA are provided in Table 4. Figures 7, 8 and 9 illustrate the evolution promoted by I12G25, I25G50 and I50G100 configurations, respectively.

Each line in Figs. 7, 8 and 9 indicates the replication of an experiment (i.e., each scenario was replicated 10 times). The initial points are different, since the population generation is random, and the best fitness found in the first generation is highly variable.

Figures 7, 8 and 9, show that the I12G25 configuration did not promote a more consistent convergence with a non-optimal result. On the other hand, the I25G50 and I50G100 settings (Figs. 8 and 9) provided higher confidence when they indicated the weights optimized for EC, since they represent a

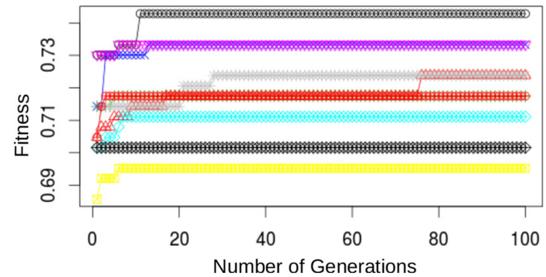


**Fig. 9** Evolution with I50G100 setting—evolution provided by 50 individuals and evolutions of 100 generations

longer period of evolution without making any advance in the area of best fitness. The period was longer than 30 generations in 80% of the cases in configuration I25G50, and longer than 70 generations in 90% of the cases in configuration I50G100. As a result, the weights indicated by the three settings of the GA were in the evaluation of the EC accuracy.

Figure 10 shows the accuracy achieved by the EC with the weights during the cross-validation evaluation phase. Despite the description of the evolution of the best fitness obtained with configuration I12G25, the weights indicated by this configuration achieved accuracy levels close to those offered by other settings (I25G50 and I50G100).

Several statistical methods have been adopted toward more reliable data analyses. The Shapiro–Wilk method was firstly used to check whether the results of the groups would fit a normal distribution. The $p$ values obtained are shown in Table 5, and confirm the normality hypothesis for the results of all the groups.

The result suggests the use of a $t$ test (a parametric method that provides higher reliability in data analysis with adjustments to the normal distribution) for checking the existence of a statistically significant difference with the results shown in Fig. 10. The $t$ test indicated that there are no statistically significant differences between all the groups of results. However, the $p$ value obtained in all pairwise comparisons was equal to one, which implies a strong similarity between the results from Ensemble and the predefined settings of the genetic algorithms. Consequently, the I12G25 configuration obtained by GA was used in the following tests.

### 6.3 Evaluation of emotion identification

Initially, we analyzed the performance of the algorithms separately through the experimental planning and evaluation technique and used $k$-fold cross-validation with $k = 10$, and $k - 1$ for training and the rest for testing. Thus, error estimation can be measured more accurately, as the mean value estimate tends to the true zero-error rate as $n$ increases, which is generally the case with small example sets. The average

**Table 4** Summary of the genetic algorithm setting

| Setting | Description | Detail/rate |
|---|---|---|
| Individual | Randomly created with minimum and maximum values | Min: 1.0 Max: 100.0 |
| Type gene | Vector of numbers (weights) | Real-valued |
| Selection | Use the fitness as a probability to be preserved in the next generation | Probabilistic |
| Mutation | Random select of point to change the value with 10% variance, respecting the probabilistic rate to perform | 0.1 |
| Crossover | Random select of point to local arithmetic of individuals, respecting the probabilistic rate to perform | 0.8 |
| Elitism | Yes. Maintain a rate of better individuals in next generation, respecting the probabilistic rate to perform | 0.05 |

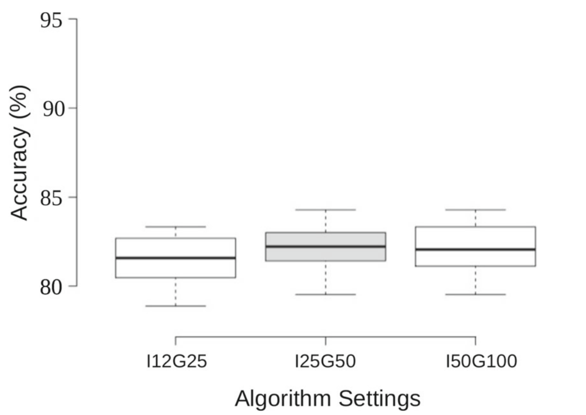


**Fig. 10** Accuracy achieved by the use of weights in different approaches assessed

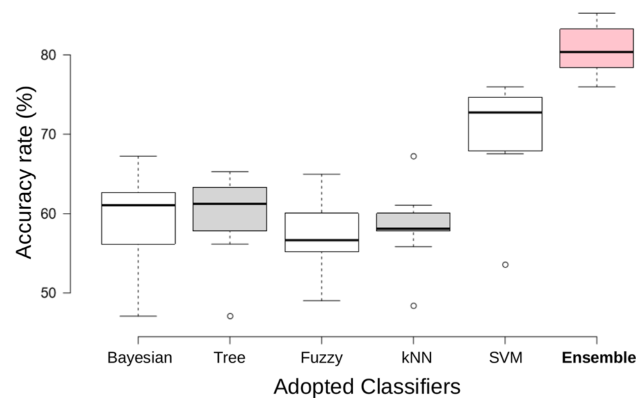


**Fig. 11** Boxplots of accuracies presented by classifiers to determine the user's emotion. Results obtained by $k$-fold cross-validation technique with $k = 10$

**Table 5** Statistic analysis of Shapiro–Wilk method approves the hypothesis of adequacy of sample sets for the normal distribution with 95% confidence interval

| Algorithm Settings | *p* value |
|---|---|
| 12 individuals and 25 generations | 0.087 |
| 25 individuals and 50 generations | 0.210 |
| 50 individuals and 100 generations | 0.192 |

*P* values greater than 0.05 indicate suitability for normal distribution

**Table 6** Median (%) accuracies and *p* values of the evaluated algorithms

| Classifiers | Accuracy (%) | *Shapiro* *p* values | Runtime (s) |
|---|---|---|---|
| N. Bayes | 59.16 | 0.115 | 0685 |
| D. Tree | 59.77 | 0.076 | 0696 |
| Fuzzy | 57.21 | 0.984 | 0862 |
| kNN | 58.41 | 0.181 | 0752 |
| SVM | 70.29 | **0.007** | 0702 |
| **EC** | **80.53** | 0.594 | 1025 |

The most relevant values shown by the tests highlighted in bold

accuracy rate for each algorithm was obtained and used in the weighting of the final decision for EC (see Sect. 5).

According to the results, EC enables a more precise classification than individual classifiers, see Fig. 11, for Boxplots for each type of classifier used. The right-hand box plot refers to the results of the EC and displays a higher median accuracy than the other classifiers (median values show in Table 6). Moreover, a smaller dispersion of the results from EC shows higher stability in its execution.

Table 6 shows the computational cost of each classifier in comparison to EC. Although the cost of EC is partially higher, this is not relevant considering the applied context that aims at interaction and classification, through a music player (this will be seen in Sect. 6.4), where classification is performed every 5 s, and the interaction with each song is performed every 3 min. We also conducted some statistical analyses to validate the results. Initially, we applied again the Shapiro–Wilk method to check their suitability for normality hypothesis, which led to parametric or nonparametric tests. Since not all the $p$ values obtained are greater than 0.05 (see Table 6), we rejected the normality hypothesis with 95% reliability, i.e., the nonparametric test was the most suitable for the next analysis.

**Table 7** *P* values of the pairwise comparison made by Wilcoxon rank sum test

| | N. Bayes | D. Tree | Fuzzy | kNN | SVM |
|---|---|---|---|---|---|
| D. Tree | 1.000 | – | – | – | – |
| Fuzzy | 1.000 | 0.783 | – | – | – |
| kNN | 1.000 | 1.000 | 1.000 | – | – |
| SVM | 0.017 | 0.017 | 0.017 | 0.017 | – |
| EC | **0.003** | **0.003** | **0.003** | **0.003** | **0.003** |

Values below 0.05 indicate statistically significant difference among the results of groups
The most relevant values shown by the tests highlighted in bold

**Table 8** Confusion matrix of the ensemble classification

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| Happy = A | **426** | 6 | 4 | 1 | 3 | 1 | 2 |
| Disgust = B | 17 | **366** | 5 | 10 | 29 | 2 | 17 |
| Fear = C | 5 | 8 | **340** | 34 | 10 | 13 | 35 |
| Neutral = D | 1 | 8 | 35 | **326** | 13 | 4 | 59 |
| Anger = E | 8 | 54 | 7 | 25 | **312** | 2 | 37 |
| Surprise = F | 1 | 1 | 13 | 2 | 3 | **423** | 2 |
| Sad = G | 7 | 16 | 24 | 47 | 31 | 4 | **316** |

The most relevant values shown by the tests highlighted in bold

The pairwise comparisons made by Wilcoxon Rank Sum test are shown in Table 7. The *p* values obtained indicate N. Bayes classifiers, D. Tree, Fuzzy and kNN do not show a statistically significant difference in the rating of emotions. However, the classification of EC (in addition to the higher accuracy shown in Fig. 11) shows a statistically significant difference with respect to other individual classifiers.

Finally, Table 8 shows the error matrix, also known as the confusion matrix. The columns of the table indicate reference data, while the lines represent the classification generated by the EC.

The model based on the Ensemble provided the best classification for all the emotions, in particular for "Joy" and "Surprise".

## 6.4 Measuring and evaluating the user's interaction: a case study with a music player

Having the aim of finding out more about the *FlexPersuade* approach, described in Sect. 6, an exploratory experiment was conducted with a group of 30 users, aged between 18 and 80. This section addresses the planning and execution and provides results of the user interaction with *FlexPersuade*. The study was approved by the Human Research Ethics Committee of the Medical School of the University of São Paulo, São Paulo, Brazil (CAAE-45081415.0.0000.0065).

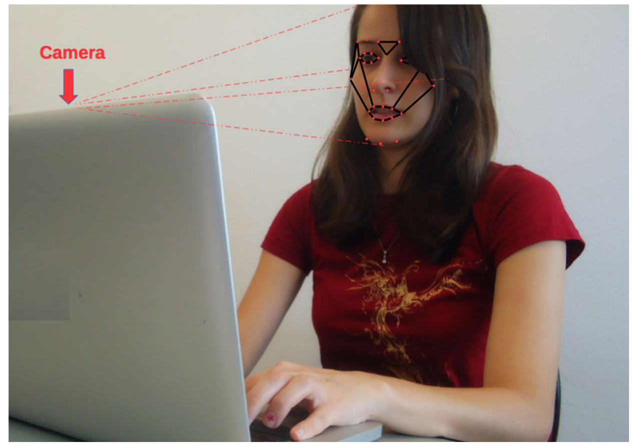


**Fig. 12** Measuring and evaluating emotions when interacting. A case study with a music player. Sensor for facial recognition assessed motor expressions; photos used of the user submitted to the experiment

### 6.4.1 Experimental planning

Prior to the evaluation of the user's interaction with *FlexPersuade*, we developed a plan to determine the user's degree of satisfaction when interacting with the music player:

- *Hypothesis* due to the diversity and increased use of computing devices, we believe *FlexPersuade* can address specific situations (i.e., emotions), regarding user's interaction with computing devices.
- *Key aim* observe and analyze the users' satisfaction with computing devices at interaction time, through assessments of their emotions (see Fig. 12).
- *Methodology* a group of 30 members, aged between 18 and 80, was invited to participate in an experiment that would analyze their interaction with flexible systems in computer devices. The users were divided into two groups. Group 1 performed the experiment with the help of *FlexPersuade* and interacted with the music player, while Group 2 performed the interaction without any interference from *FlexPersuade*, i.e., the songs were played randomly. Table 9 shows the division of the groups.
  The experiments were divided for determining the degree of user's satisfaction when interacting with *FlexPersuade*. The users were required to use the notebook according to their preferences, such as using social networks, reading news and writing an e-mail, while listening to music from the system.
- *Supporting material* the documentary material prepared for the exploratory experiment involved an Informed Consent Term, a Term for Authorized Capture and Display of Images, and a Profile Survey Questionnaire. The Informed Consent Term clarified to the participants the purpose of the research, its scientific character and the

**Table 9** Division of groups by age for the experiments

| | 18 to 30 | 31 to 50 | 50 years old |
|---|---|---|---|
| Group 1 | 5 | 5 | 5 |
| Group 2 | 5 | 5 | 5 |

**Table 10** Confusion matrix of the classification of images of users subjected to the Ensemble test

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| Happy = A | **30** | 0 | 0 | 0 | 0 | 0 | 0 |
| Disgust = B | 2 | **21** | 0 | 0 | 4 | 0 | 3 |
| Fear = C | 0 | 0 | **21** | 4 | 1 | 2 | 2 |
| Neutral = D | 0 | 0 | 2 | **25** | 0 | 0 | 3 |
| Anger = E | 0 | 6 | 0 | 1 | **19** | 0 | 4 |
| Surprise = F | 0 | 0 | 1 | 0 | 0 | **29** | 0 |
| Sad = G | 0 | 1 | 1 | 4 | 2 | 0 | **22** |

The most relevant values shown by the tests highlighted in bold

voluntary nature of their participation. The Term for the Authorized Capture and Display of Images, Sounds and Names explained to the participants the experiment would involve a "capture" of images and the data would be used for scientific purposes.

– *Devices* the users interacted on an individual basis. During a 60 min interaction, the images were captured by a camera on the computer. An ASUS® Intel® Core i5 laptop, RAM 4 GB and HD 500 GB was used.

### 6.4.2 Discussion of the results

The experiments were performed in two steps. The first step focused on the analysis of the performance related to the accuracy obtained from the classification images with images that had not been used to train the classification model. The second step revealed the user's satisfaction obtained from the music player which was regarded as a flexible solution.

Table 10 shows the first evaluation and the confusion matrix achieved by the EC, based on the weighted average. 30 images of each emotion per participant were used, resulting in 210 images. They were not used for training purposes, i.e., they were unknown and achieved 79.04% accuracy levels.

The second evaluation focused on the group's satisfaction, represented in the graph in Fig. 13 showing the data of the Participant Observation Form. The members performed a 60 min test: every 15 min they were asked to state their satisfaction on a scale from 0 to 10, using the *FlexPersuade* system. The satisfaction of users who interacted with *FlexPersuade* (Group 1) was higher than that of the users who did not interact with the music player through *FlexPersuade* (Group 2).

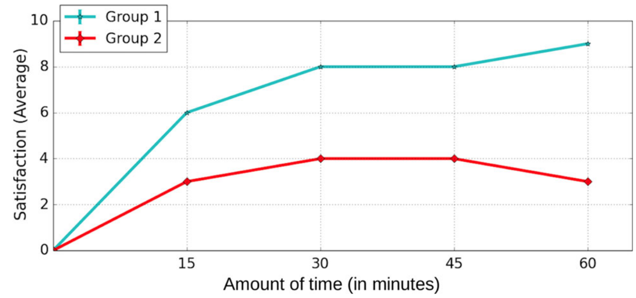


**Fig. 13** Satisfaction of user's interaction (with and without *FlexPersuade*)

Based on the experiments, we can infer that *FlexPersuade* largely fulfilled its purpose, i.e., identification of users' emotional state and suggestion of music that could change (i.e., persuade) their emotion, around 60% more when the *FlexPersuade* interaction is not performed.

## 7 Concluding remarks

Several methods, techniques and tools in the field of HCI support the evaluation of emotional responses, including the capture, analysis and classification of users' emotional responses. We investigated and implemented a smarter model by exploiting the EC classification concept, so that the error generated by each individual classifier could be reduced through the combination of other components. Based on our generic proposal, we presented an application of music player for the detection of emotions and interaction based on the user's face.

The results showed that the Ensemble classification achieves a high degree of reliability with 83% average accuracy, higher than other techniques used to classify the user's emotions. The characteristics of emotions, such as happiness, surprise and anger, facilitate their identification, while other emotions can be difficult to interpret because of the confusing facial expressions. The evaluation of the approach was based on genetic algorithm for the attribution of weights to each algorithm for the final decision of Ensemble.

When it was tested with users, there was a significant satisfaction in the use of *FlexPersuade*. A satisfaction increase of 60% was observed when the music player case study was included. We argue that a generic and adaptive model can better meet user's ways of interaction, since each of them has a particular interaction and an intelligent and generic model can learn and adapt to various forms of interaction of the user.

### 7.1 Applicability and future directions

*Flexpersuade* is a model developed for systems of recognition and emotional interaction through the face and can be used/applied in several contexts for assisting users, in addition to the music player demonstrated in this study.

In education, *Flexpersuade* can improve the students' experience in a learning environment, assist teachers in teaching methodologies, as well as virtual training programs. Ambient-Assisted Living or Health Smart Homes can be applied to the monitoring of emotional aspects of individuals, especially those with disabilities or the elderly, and assist in decision-making processes or issue alerts for medical staff and/or family members. Moreover, it can help in the diagnoses of diseases, such as schizophrenia, depression, autism and bipolar disorder, which may be caused by emotions, lack of emotional expressions or instability of emotions.

Despite the numerous applications and good results of *Flexpersuade*, its accuracy can be improved. Therefore, we intend to explore:

- other approaches that identify emotions, e.g., Trend Component Behavioral—Social Networks and voice and speech analysis;
- results of the combination of multiple sensors for identifying emotions;
- interaction/persuasion strategies that best meet users' preferences, and;
- use of *Flexpesuade* in real environments, such as education, Ambient-Assisted Living and health area to deploy both recognition and use of emotional aspects toward helping users.

**Acknowledgements** Authors would like to thank the support from the University of Groningen and Professor Michael Biehl of the University of Groningen.

**Funding** This study was funded by Foundation for Research Support of the State of Sao Paulo—FAPESP (Grant Nos. 2016/14267-7 and 2017/21054-2).

## Compliance with ethical standards

**Conflict of interest** Author Leandro Y. Mano declares that he has no conflict of interest. Author Bruno S. Faiçal declares that he has no conflict of interest. Author Vinícius P. Gonçalves declares that he has no conflict of interest. Author Gustavo Pessin declares that he has no conflict of interest. Author Pedro H. Gomes declares that he has no conflict of interest. Author Andrç C. P. L. F. de Carvalho declares that he has no conflict of interest. Author Jó Ueyama declares that he has no conflict of interest.

**Ethical approval** All procedures performed in studies involving human participants were in accordance with the Human Research Ethics Committee of the Medical School of the University of São Paulo, São Paulo, Brazil (CAAE-45081415.0.0000.0065). This article does not contain any studies with animals performed by any of the authors.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

## References

Bailenson JN, Pontikakis ED, Mauss IB, Gross JJ et al (2008) Real-time classification of evoked emotions using facial feature tracking and physiological responses. Int J Hum Comput Stud 66(5):303–317. https://doi.org/10.1016/j.ijhcs.2007.10.011

Bouckaert RR, Frank E, Hall M et al (2013) WEKA Manual for Version 3-7-8

Carvalho A, Faceli K, Lorena AC, Gama J (2011) Inteligência Artificial-uma abordagem de aprendizado de máquina. LTC, Rio de Janeiro

Coulson G, Blair G, Grace P et al (2008) A generic component model for building systems software. ACM Trans Comput Syst 26(1):1–42. https://doi.org/10.1145/1328671.1328672

Darwin C, Wallace A (1858) On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. J Proc Linn Soc Lond 3(9):45–62. https://doi.org/10.1111/j.1096-3642.1858.tb02500.x

Darwin C, Adler MJ, Hutchins RM (1872) The origin of species by means of natural selection. J. Murray, London

Duda RO, Hart PE, Stork DG (2012) Pattern classification. Wiley, New York

Ebner NC, Riediger M, Lindenberger U (2010) FACES–a database of facial expressions in young, middle-aged, and older women and men: development and validation. Behav Res Methods 42(1):351–362. https://doi.org/10.3758/BRM.42.1.351

Ekman P, Friesen WV, O'Sullivan et al (1987) Universals and cultural differences in the judgments of facial expressions of emotion. J Personal Soc Psychol 53(4):712–717. https://doi.org/10.1037/0022-3514.53.4.712

Gonçalves VP, Costa EP, Valejo A, Geraldo Filho PR, Johnson TM, Pessin G, Ueyama J (2017a) Enhancing intelligence in multimodal emotion assessments. Appl Intell 46(2):470–486. https://doi.org/10.1007/s10489-016-0842-7

Gonçalves VP, Giancristofaro GT, Geraldo Filho PR, Johnson T, Carvalho V, Pessin G, Ueyama J (2017b) Assessing users' emotion at interaction time: a multimodal approach with multiple sensors. Soft Comput 21(18):5309–5323. https://doi.org/10.1007/s00500-016-2115-0

Gonçalves VP, de Almeida Neris VP, Seraphini S, Dias TC, Pessin G, Johnson T, Ueyama J (2017c) Providing adaptive smartphone interfaces targeted at elderly people: an approach that takes into account diversity among the elderly. Univers Access Inf Soc 16(1):129–149. https://doi.org/10.1007/s10209-015-0429-9

Kalsi KS, Rai P (2017) A classification of emotion and gender using approximation image Gabor local binary pattern. In: 7th International conference on cloud computing, data science and engineering-confluence. IEEE Explore, pp 623–628. https://doi.org/10.1109/CONFLUENCE.2017.7943227

Klein J, Moon Y, Picard RW (2002) This computer responds to user frustration: theory, design, and results. Interact Comput 14(2):119–140. https://doi.org/10.1016/S0953-5438(01)00053-4

Langner O, Dotsch R, Bijlstra G et al (2010) Presentation and validation of the Radboud Faces database. Cognit Emot 24(8):1377–1388. https://doi.org/10.1080/02699930903485076

Lee YH, Ahn H, Cho HJ, Lee JH (2015) Recognition of facial emotion through face analysis based on quadratic bezier curves. Indian J Sci Technol 8(35):634–641. https://doi.org/10.17485/ijst/2015/v8i35/86092

Libralon GL, Romero RAF (2014) Mapping of facial elements for emotion analysis. In: Brazilian conference on intelligent systems (BRACIS). IEEE Xplore, pp 222–227. https://doi.org/10.1109/BRACIS.2014.48

Lichtenstein A, Oehme A, Kupschick S, Jürgensohn T, Peter C, Beale R (2008) Comparing two emotion models for deriving affective states from physiological data. In: Affect and emotion in human–

computer interaction: from theory to applications. Springer, Berlin, pp 35–50. https://doi.org/10.1007/978-3-540-85099-1_4

LiKamWa R, Liu Y, Lane ND, Zhong L (2013) MoodScope: building a mood sensor from smartphone usage patterns. In: Proceeding of the 11th annual international conference on Mobile systems, applications, and services (MobiSys '13). ACM, New York, pp 389–402. https://doi.org/10.1145/2462456.2464449

Lucey P, Cohn JF, Kanade T, Saragih J et al (2010) The extended Cohn–Kanade dataset (ck+): a complete dataset for action unit and emotion-specified expression. In: IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW), 2010, IEEE Explore, pp 94–101. http://doi.acm.org/10.1109/CVPRW.2010.5543262

Mahlke S, Minge M (2008) Consideration of multiple components of emotions in human–technology interaction. In: Peter C, Beale R (eds) Affect and emotion in human–computer interaction. Lecture notes in computer science. Springer, Berlin, pp 51–62. https://doi.org/10.1007/978-3-540-85099-1_5

Mano, LY, Giancristofaro GT, Faiçal BS et al (2015) Exploiting the use of ensemble classifiers to enhance the precision of user's emotion classification. In: Proceedings of the 16th international conference on engineering applications of neural networks (INNS). ACM, New York, pp 1–7. https://doi.org/10.1145/2797143.2797165

Mano, LY (2018) Emotional condition in the Health Smart Homes environment: emotion recognition using ensemble of classifiers. In: 2018 Innovations in intelligent systems and applications (INISTA). IEEE Explore, pp 1–8. http://doi.acm.org/10.1109/INISTA.2018.8466318

Mano LY, Faiçal BS, Nakamura LHV, Gomes PH et al (2016a) Exploiting IoT technologies for enhancing Health Smart Homes through patient identification and emotion recognition. Comput Commun 89:178–190. https://doi.org/10.1016/j.comcom.2016.03.010

Mano LY, Vasconcelos E, Ueyama J (2016b) Identifying emotions in speech patterns: adopted approach and obtained results. IEEE Latin Am Trans 14(12):4775–4780. https://doi.org/10.1109/TLA.2016.7817010

Mano LY, Sawada LO, Ueyama J (2016c) Abordagem para a interação afetiva: um estudo de caso com player de musica. J Adv Theor Appl Inf 2(2):47–54. https://doi.org/10.26729/jadi.v2i2.2108

Mena CJP, Velho L, Junior RMC (2011) 3D human face reconstruction using principal components spaces. In: Proceedings of WTD SIBGRAPI conference on graphics, patterns and images, Brazilian computer society. pp 1–6

Peter C, Urban B (2012) Emotion in human–computer interaction. In: Dill J, Earnshaw R, Kasik D, Vince J, Wong P (eds) Expanding the frontiers of visual analytics and visualization. Springer, London, pp 239–262. https://doi.org/10.1007/978-1-4471-2804-5_14

Russell JA (1980) A circumplex model of affect. J Personal Soc Psychol 39(6):1161–1178. https://doi.org/10.1037/h0077714

Russell JA (1994) Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. Psychol Bull 115(1):102–141. https://doi.org/10.1037/0033-2909.115.1.102

Saragih JM, Lucey S, Cohn JF (2011) Deformable model fitting by regularized landmark mean-shift. Int J Comput Vis 91(2):200–215. https://doi.org/10.1007/s11263-010-0380-4

Scherer KR (2005) What are emotions? And how can they be measured? Soc Sci Inf 44(4):695–729. https://doi.org/10.1177/0539018405058216

Schuller B, Reiter S, Muller R et al (2005) Speaker independent speech emotion recognition by ensemble classification. In: IEEE International conference on multimedia and expo, 2005, ICME 2005. IEEE Explore, pp 864–867. https://doi.org/10.1109/ICME.2005.1521560

Shubham G, Kusum D (2019a) A novel random walk grey wolf optimizer. Swarm Evolut Comput 44:101–112. https://doi.org/10.1016/j.swevo.2018.01.001

Shubham G, Kusum D (2019b) Improved sine cosine algorithm with crossover scheme for global optimization. Knowl Based Syst 165:374–406. https://doi.org/10.1016/j.knosys.2018.12.008

Shubham G, Kusum D (2019c) A hybrid self-adaptive sine cosine algorithm with opposition based learning. Expert Syst Appl 119:210–230. https://doi.org/10.1016/j.eswa.2018.10.050

Tian Y, Kanade T, Cohn JF (2011) Facial expression recognition. In: Li S, Jain A (eds) Handbook of face recognition. Springer, London, pp 487–519. https://doi.org/10.1007/978-0-85729-932-1_19

Zhou F, Qu X, Helander MG, Jiao JR (2011) Affect prediction from physiological measures via visual stimuli. Int J Hum Comput Stud 69(12):801–819. https://doi.org/10.1016/j.ijhcs.2011.07.005