CrossMark

A fuzzy logic approach to reliable real-time recognition of facial emotions

Kiavash Bahreini [1] · Wim van der Vegt [1] · Wim Westera [1]

Abstract

This paper represents our newly developed software for emotion recognition from facial expressions. Besides allowing emotion recognition from image files and recorded video files, it uses webcam data to provide real-time, continuous, and unobtrusive facial emotional expressions. It uses FURIA algorithm for unordered fuzzy rule induction to offer timely and appropriate feedback based on learners' facial expressions. The main objective of this study was first to validate the use of webcam data for a real-time and accurate analysis of facial expressions in e-learning environments. Second, transform these facial expressions to detected emotional states using the FURIA algorithm. We measured the performance of the software with ten participants, provided them with the same computer-based tasks, requested them a hundred times to mimic specific facial expressions, and recorded all sessions on video. We used the recorded video files to feed our newly developed software. We then used two experts' opinions to annotate and rate participants' recorded behaviours and to validate the software's results. The software provides accurate and reliable results with the overall accuracy of 83.2%, which is comparable to the recognition by humans. This study will help to increase the quality of e-learning.

Keywords Emotion recognition · Affective computing · Software development · Statistical data analysis · Fuzzy logic · Webcam, E-learning

1 Introduction

1.1 Emotion and e-learning

Emotions are a significant influential factor in the process of learning [58]. Current instructional methods for online learning increasingly address emotional dimensions by

---

✉ Kiavash Bahreini
    kiavash.bahreini@ou.nl; kiavash.bahreini@yahoo.com; kbsocialnetwork@gmail.com

[1]   Welten Institute, Research Centre for Learning, Teaching and Technology, Faculty of Psychology and
      Educational Sciences, Open University of the Netherlands, Heerlen, Netherlands

accommodating challenges, excitement, ownership, and responsibility among other things in the learning environment [25, 80]. Educational games [18] would be the case in point: offering a challenging and dynamic learning setting that effortlessly combines emotion and cognition [78]. While online learning has expanded radically over the past years, there is a renewed interest in adaptive methods and personalization that adjust the instruction and support explicitly to the learners' mental states and requirements. Such personalization is conventionally based on producing and maintaining a model of the learner, which is mainly based on individual characteristics and validated performances [13, 14]. Emotion has systematically been ignored as a learner model variable because it was hard, if not impossible to detect. Now that technology is about to be capable of automatically recognising the learners' emotional states, learner models could readily include emotions and thereby improve the quality of personalization.

Emotion recognition in e-learning environments could, while obviously taking into account issues of ethics and privacy, propose a valuable source for improving the quality of learning [11]. Responses based on emotional states [35] could enhance learners' understanding of their own performance [10].

Also, educational game development [79] can take advantage of emotional data of learners [4] to optimise experiences and the flow of events. In this study, we offer an accurate and reliable technology for emotion recognition that can be easily applied in digital educational games and other e-learning environments.

1.2 Approaches to emotion recognition

Technologies related to emotion recognition date back to the early 1900s. Different tests of blood pressure were used for lie recognition during the questioning of criminal suspects at that time [53]. Although occasionally lie detectors are admitted as evidence in court, however, they are generally considered unreliable. Over the years, there have been several improvements in the accuracy of emotion recognition software. Bettadapura [12] reports accuracies for existing emotion recognition solutions ranging from 55% to 98% since 2001. Basically, six different approaches to emotion recognition are available, ranging from 1) using facial expressions [7], 2) speech and vocal intonations [8], 3) physiological signals [65], 4) body gesture and pose [59], 5) text [66], and 6) a combination of the two or more of these approaches [9, 60, 64].

Facial expressions provide the most informative data for computer awareness of emotions [64]. However, software applications that use facial expressions have a number of restrictions that mostly limit their accuracy and applicability. Usually, they can only manage a small set of expressions from a frontal view of faces without facial hair and glasses, and they require good and stable lighting conditions. Also, most software applications cannot be used in real-time, but require extensive post-processing for the analysis of videos and images [57].

Emotion recognition can also be provided based on the audio and vocal intonations in recorded speech [19]. However, the analysis of vocal intonations produces less accurate emotion recognition results than facial expressions approaches. Vocal intonation analysis can only manage a subset of basic emotions from the recorded audio files or from the speech streams that come from microphones. These require post-processing through various speech analysis methods [75].

Physiological sensors allow for capturing a variety of physiological responses such as body temperature, heart rate, blood volume, and skin conductance of an individual [42]. These

sensors are sometimes offered in the form of wearable devices [69]. They sometimes are used to study experienced emotions of learners in schools [3], and as an add-on to intelligent tutoring systems [31]. Although such technologies show promising results in emotion recognition, it is scarcely applied, because it is obtrusive to learners and requires expensive and dedicated equipment [46].

Body movements and gestures are an additional source of emotion recognition [40, 76]. D'Mello and Graesser report that there are significant relationships between emotion and body movements. The combination of posture and conversational dialogue systems reveals a modest amount of redundancy among them [17]. Kyriakos and colleagues developed a real-time dynamic hand gesture and posture recognition system for the formation of a visual dictionary by merging hand postures and dynamic gestures [43]. Recently, the video games industry has introduced sensory devices as a commodity, mainly meant for interaction control in entertainment games [61]. Already in 2005, Nintendo introduced its Wii game console, with a movement and gesture recognition sensor. Likewise, Microsoft introduced in 2010 its Kinect to provide optical sensor technology for body recognition and motion tracking [84]. Although both Wii and Kinect have greatly enhanced gaming interaction modes by supporting capturing gestures and bodily movements, they are not capable of extracting the users' emotions [68]. Emotion recognition through text or speech analysis is applied to a set of words in a specific language [48]. Such analysis is called sentiment analysis and uses natural language processing techniques for extracting the affective state represented in the text and thereby the affective state and attitude of the author of a text [81]. Dependency on a specific language is the main obstacle in developing a worldwide software application for recognizing emotions from the text [48]. Another obstacle is where the speakers or authors do not necessarily express their own emotions, but describe somebody else's emotion [51] or sometimes they do not express the emotion in a sentence explicitly [16]. Some studies reported that these issues could be solved using semantic technologies (see for example [16]). Such technology adds some metadata over the textual data and encodes the meaning of the text [5, 6].

The accuracy of emotion recognition can be greatly improved by combining two or more of the previous approaches. Jaimes and Sebe [34] have shown improved performance by combining visual and audio information. They showed that the multimodal data fusion could rise to accuracy levels from 72% up to 85% if the following conditions are met: 1) clean audio-visual input, such as noise free dataset, closed and fixed microphone, and non-occluded portraits, 2) with actors' performances, 3) who speak single words, and 4) who display exaggerated facial expressions of the six basic emotions (happiness, sadness, surprise, fear, disgust, and anger) (cf. [20]).

1.3 Emotion recognition in e-learning

Notwithstanding the limitations of emotion recognition described above, topical hardware developments on regular computer equipment [23] would now enable emotion recognition at a larger scale [7]. A typical example would be the use of common webcams for emotion recognition from facial expressions [7, 52]. It has been suggested that e-learning applications can benefit from such emotion recognition devices for more natural interactions [71] because they collect data of learners continuously and unobtrusively [7, 8]. For a long time approaches for collecting emotional data of learners have been either obtrusive or discontinuous [67]. For

example, physiological sensors and questionnaires can fundamentally hinder the learning process [24] and they are not convenient or appropriate to use in e-learning environments [63]. Using webcam-based approaches to emotion recognition would overcome these problems. Various problems have been reported though. Emotion recognition from facial expressions could not be detected in real-time from the frontal view of faces [82]. Intensive post-processing is often needed to analyse recorded video files or stored image files of learners [38]. Occasional solutions for real-time recognition of facial emotions produced low accuracies that are not comparable to emotion recognition by humans [7]. It has been difficult to accurately detect faces and facial emotions when beard, glasses, hair over face, wounds, or other objects cover any parts of the face [47]. Moreover, recognition is hampered when disturbing light shines directly into the face of the learner.

1.4 Several techniques for classification of facial emotion recognition

Researchers have proposed many methods for recognising and classifying emotions from facial expressions. Prior studies show that there are many different techniques to distinguish facial expressions. However, we only report eight of the most notable methods in this study: 1) pixel-based recognition [77], 2) local binary pattern [50], 3) wavelet transform [36], 4) discrete cosine transform [37], 5) Gabor filter [56], 6) edge and skin detection [32], 7) facial contour [26], and 8) fuzzy logic model [22]. Each of these studies has shown that the facial emotion recognition can accomplish an average level of success, but the performance is less than human judgement. These studies have shown that the accuracy of automatic facial emotion recognition classifications remains challenging because of the inconstancy, complexity, hard to implement, and inappropriate tracking of facial features in real-time or recorded video streams. As a result, we introduce a new approach using fuzzy logic rules to generate better, faster, more accurate, and more reliable results.

Recent studies have shown that researchers can recognise and classify the facial emotional expressions more appropriately. For example, Ali and her colleagues [1] have proposed an application of nonlinear and non-stationary data analysis techniques named, Empirical Mode Decomposition (EMD) that can classify the facial emotions with better accuracy compared to the stated methods. They have used static images as input to their application and have extracted facial features accordingly. They have applied ANOVA test as the statistical data analysis technique to obtain the facial features that were statistically significant. Then they have sent the facial features into the algorithms such as K-NN and SVM for classification of seven categories of facial emotions. In another study, Gunes and Pantic [27] have considered Russell's method for circular configuration called Circumflex of Affect [62]. In this method, every primary emotion illustrates a bipolar entity as an element of the similarly emotional continuum. The suggested polar are valence (pleasant versus unpleasant), and arousal (relaxed versus aroused). The recommended emotional space comprises four quadrants: high arousal positive, low arousal positive, high arousal negative, and low arousal negative. Consequently, it is likely to represent every emotion by its valence and arousal. Moreover, they have investigated automatic, dimensional, and continuous emotion recognition using visual, audio, physical, and brainwave methods on their study. Their findings revealed that representing emotions continuously is not a small problem to ignore and to handle easily. In another study, Anisetti and his colleagues [2] proposed a semi-supervised fuzzy facial emotional classification system based on Russell's circumplex model. Their proposed system works only on face related features

classified with the Facial Action Coding System (FACS). They have extracted facial emotional expressions from streaming videos. To evaluate the quality of their system, they have used the Cohn-Kanade database and the MMI Database to apply Russell's space mapping. To make Russell's axes classification, they have created an emotional inference space and have mapped action units to axes values. Then, they have exploited some well-defined rules from this mapping. Although they proposed this novel system, however, the system requires expert tuning to guarantee context awareness. They have concluded that researchers should further investigate on tuning the system to obtain better outputs for facial emotional classification in the complex scenarios.

1.5 Starting point

In this paper, we present a new methodology of webcam-based emotion recognition, along with a full technical implementation that was used for its validation. The approach is based on fuzzy logic, using unordered fuzzy rule induction (FURIA algorithm; [29]). Compared to the statistically data analysis approach proposed by Ali and her colleagues [1], our fuzzy logic approach uses the supervised machine learning method to provide more favourable output because fuzzy logic rules can be easily generated from a dataset of recorded emotions, while alternative machine learning approaches, such as neural networks, Bayesian networks, and decision trees would require extensive implementation. Moreover, our approach can use single image files, recorded video files, and live webcam streams to propose an accurate recognition of facial expressions compared to the approach suggested by Ali and her colleagues that can only use single image files [1]. We follow the emotion classification approach of Ekman and Friesen [20], which has been frequently used over the past decades for classifying the six basic emotions: happiness, sadness, surprise, fear, disgust, and anger.

We do not follow Russell's method for Circumflex of Affect used by Gunes and Pantic [27]. Therefore we do not calculate bipolar entities such as high arousal positive, low arousal positive, high arousal negative, and low arousal negative in our approach. Instead, we use extracted facial features by tracking a human face in real-time and classify facial emotions. Moreover, compared to the semi-supervised fuzzy facial emotional classification system based on Russell's circumplex model proposed by Anisetti and his colleagues [2], we recommend a new approach that can classify emotions based on the FURIA fuzzy rules using the supervised machine learning technique. Our rules do not need to be produced based on the emotional inference space and mapped action units to axes values. Instead, we generate our rules based on the cosine values of the most significant triangles created based on the most significant facial feature points. We will describe our approach in the coming sections.

Although similar to our previous approach [7], which used Principal Component Analysis, the fuzzy logic-based approach produces better, more accurate, and more reliable results. To allow maximal portability the software is implemented as a RAGE-compliant software component [73, 74]: the RAGE software architecture omits dependencies on platforms and operating systems and accommodates the easy reuse and integration of software in a variety of video game engines. In the rest of this paper, we first describe creating a facial emotion database and the functionalities of our software. Thereafter, we explain the validation method used in this study, discuss the results of this study, and provide suggestions for future work, respectively.

2 Database, fuzzy rules, and software

2.1 Creating a database of the facial emotions

We started from an existing database, the Cohn-Kanade AU-coded expression extended database (CK+) as the reference to this study [49]. This database is used for automatic facial image analysis and includes an annotated set of human's facial images, including validated emotion labels for each image. Based on this, we then created a database of emotions including the rotated images of each subject. We then created cosine values of facial landmarks for training and testing purposes. This database then was used to deduce fuzzy rules. To this end, we developed a small software application that used DLIB [39], which is a widely used C++ toolkit including machine learning tools and algorithms. After loading images and their related emotion labels from the CK+ database, face recognition and face tracking functionalities from DLIB were used and extended to develop facial emotion classification functionality. From each image, we extracted 68 facial landmarks and made 54 vertices for 18 relevant triangles using every three important landmarks in our database. For example, two important triangles with 6 vertices are the triangles between eyebrows and eyes (see Fig. 1, facial landmarks 17, 36, and 39 & 22, 42, and 45). We then calculated the cosine values (54 values) of all vertices in all triangles. Next, we stored all the cosine values along with the related emotion labels of each image of the CK+ database in our database in the form of a WEKA attribute-relation file format (arff) [70]. WEKA is a tool that provides a number of machine learning algorithms for data mining tasks. The arff file is a textual database that defines a list of instances sharing a set of attributes: each instance is represented with 55 attributes called Cosine0, Cosine1, …, Cosine53 and Emotion, respectively. By loading the database in WEKA 37 so-called FURIA fuzzy rules (see appendix 1) could be generated, allowing us to automatically detect and classify emotions from facial expressions. FURIA is a fuzzy rule-based classification method, which offers simple and comprehensible rule sets [29]. WEKA does not provide the FURIA rule-based classifier algorithm as default; therefore users must add this classifier algorithm to the existing classifiers. Users can use the package manager of the WEKA tool in the WEKA GUI Chooser to install FURIA before they run the WEKA Explorer application. When users added the FURIA classifier to the list of the WEKA classifiers, they can run the FURIA classifier and produce the FURIA fuzzy rules. The mechanism of fuzzy rules is briefly explained in the next section.
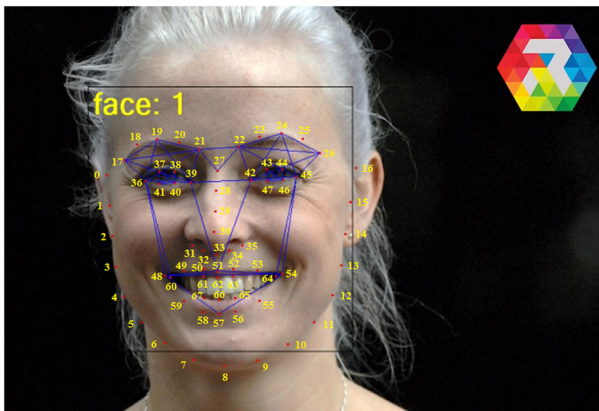

Fig. 1 A detected face, facial landmarks, the vertices, and the relevant triangles of the face

2.2 Fuzzy rules

A fuzzy rule is obtained by replacing binary logic intervals with fuzzy intervals. For example, a binary interval would be represented as step or block function (with a discrete value of 1 ("true") if the parameter under consideration is inside the interval, and 0 ("false") elsewhere). A fuzzy rule, however, could be shaped as a trapezium, allowing for "fuzzy" truth-values between 0 and 1 (Fig. 2). This can be formalised as follows: the trapezoidal membership function for a fuzzy set F on the universe of discourse X is defined as $\mu F : X \rightarrow [0,1]$, where each element of X is mapped into a value between 0 and 1. This function is defined by four parameters [29]: a lower limit LL, an upper limit UL, a lower support limit LSL, and an upper support limit USL, where $LL < UL < LSL < USL$:

$$\mu F : X =$$

$$
\begin{array}{ll}
0, & (X <= LL) \text{ or } (X > USL) \\
(X-LL)/(UL-LL), & LL <= X <= UL \\
1, & UL <= X <= LSL \\
(USL-X)/(USL-LSL), & LSL <= X <= USL
\end{array}
$$

The four parameters of the trapezoidal member function are indicated on the horizontal axis in Fig. 2.

We have generated 37 FURIA fuzzy rules in this study. Appendix 1 presents all the rules. As an example, we explain one of our generated FURIA fuzzy rules (rule number 10) to show how the emotion recognition logic is expressed. Fuzzy rule number 10 reads as follows:

^{B}(Cosine1 in [−inf, −inf, 6.82602, 7.03498]) and (Cosine15 in [14.5889, 15.0512, inf, inf])

$=>$ Emotions = Sad (CF = 0.53)$^{\wedge}$.

The antecedence of the rule includes two trapezoidal conditions. The arguments between brackets represent the 4 trapezoidal parameters. As inf indicates infinity, both trapeziums in this example are degenerate. The overall rule can be interpreted as:

(1)   Cosine1 in [−inf, −inf, 6.82602, 7.03498]: This expression is completely valid for Cosine1 <= 6.82602, invalid for Cosine1 > 7.03498, and partially valid in-between
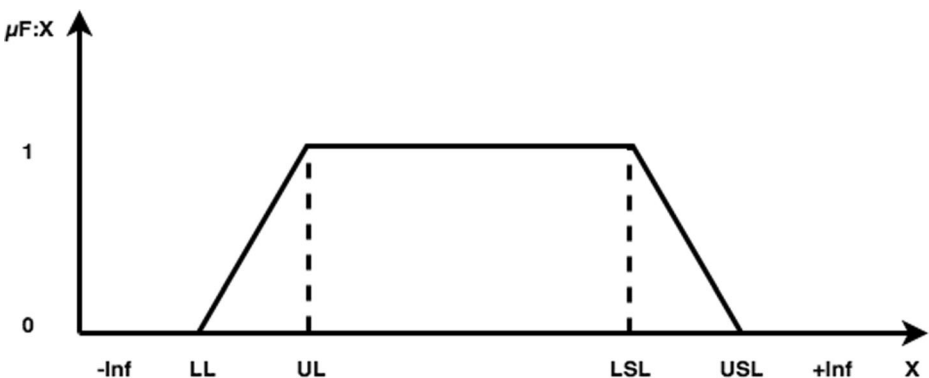

Fig. 2 Binary logic and the trapezoidal member function of a fuzzy interval

(2)   Cosine15 in [14.5889, 15.0512, inf, inf]: it is invalid for Cosine15 < 14.5889, completely
       valid for Cosine15 >= 15.0512, and partially valid in-between.
(3)   This rule means that if the aforementioned conditions are met, then the emotion will be
       considered to be "sad".

2.3 Implementation of emotion recognition from facial expressions

The software was developed in accordance with the RAGE client asset architecture [73, 74],
which prohibits direct access to the operating system and hardware. As a result, the software
accepts raw image data that can originate from various sources such as pictures or screenshots and
frames from either pre-recorded video or live webcam streams making is very versatile in its
application. The process of emotion recognition starts with face recognition. This is done using
DLIB [39], which provides functionality for real-time tracking for not losing the face. It also
provides a sufficient set of 68 landmarks, which reflect the significant positions on the individual's
face, which are dynamically updated. Once the 68 facial landmarks of a face are extracted, we
overlay 18 relevant triangles on the face. We then calculate 54 cosine values of all vertices of the
triangles. Next, the fuzzy rules come into play: all 54-cosine values are passed into the rules set to
extract and classify the expressed emotion. Figure 3 represents the software with 3 detected faces.

3 Validation method

The validation of the approach is arranged by asking test persons to express a series of
emotions and compare the judgements by the fuzzy logic approach with the judgements made
by experts. For this, we used the recorded video files of test persons from a previous study [7]
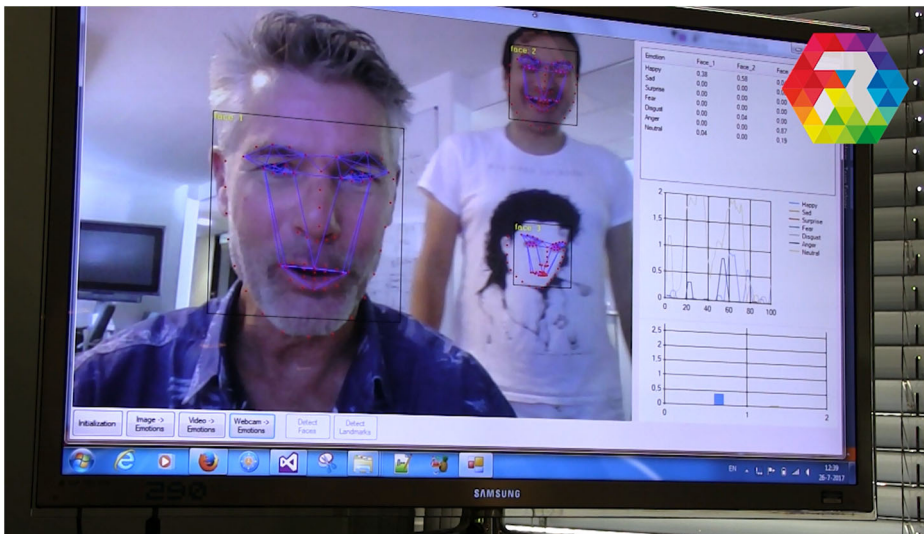to feed into the fuzzy logic system. The whole procedure is described below.

Fig. 3 The software with 3 detected faces. Two faces are faces of real persons in real-time and the third face is a
drawing face of Michael Jackson printed on a T-Shirt

3.1 Participants

We have sent an email out to employees from the Open University of the Netherlands to recruit the participants for this study. The e-mail mentioned the estimated time investment of 20 min for enrolling in the study. Activities entailed the active expression of a series of facial emotions. No specific background knowledge was requested. Ten participants, all employees from Open University of the Netherlands (8 male, 2 female; age M = 42, SD = 11) volunteered to participate in the study. Altogether, this small number of participants was sufficient for generating a dataset of 1000 facial expressions. By signing an agreement form, the participants allowed us to capture their facial expressions and to use their data anonymously for future research. We assured the participants that their raw data would not be available to the public, would not be used for commercial or similar purposes, and would not be available to third parties. Participants were told that participation in the study might help them to become more aware of their emotions while they were communicating through a webcam with our software.

3.2 Tasks

Five consecutive tasks were given to the participants. Participants were asked to expose six basic facial expressions as well as the neutral one. Totally, facial expressions were requested one hundred times, uniformly distributed over the six emotions and the neutral emotion. Each of tasks serves a different purpose. The first task was meant to calibrate the user's facial expressions. In the second task, participants were asked to mimic a pre-set emotion that was presented in an image shown to them. There were 35 images presented subsequently through PowerPoint slides; the participant scrolled through the slides. Each image illustrated a single emotion. All six basic facial expressions and the neutral one were five times present with the following order: happy, sad, surprise, fear, disgust, anger, neutral, happy, etcetera. In the third task, participants were requested to mimic the seven facial expressions twice: first, through slides that each presented the keyword of the requested emotion and second, through slides that each presented the keyword and the picture of the requested emotion with the following order: anger, disgust, fear, happy, neutral, sad, and surprise. The fourth task presented 14 slides with the text transcript (both sender and receiver) taken from a good-news conversation. The text transcript also included instructions what facial expression should accompany the current text-slide. Here, participants were requested to read and speak aloud the sender text of the 'slides' from the transcript and show the accompanying facial expression. The fifth task with 30 slides was similar to task 4, but in this case, the text transcript was taken from a bad-news conversation. The transcripts and instructions for tasks 4 and 5 were taken from an existing Open University of the Netherlands (OUNL) training course [45] and a communication book (Van der [72]).

3.3 Hardware and software

Participants performed individually on a single computer. The computer screen was separated into two panes, left and right. The tasks and the PowerPoint file were presented in the right pane, while the participants could read in the left pane how the software classified their facial expressions. An integrated webcam and a 1080HD external camera were used to capture and record the emotions of the participants as well as their interactions with mouse and keyboard

on the computer screen. The integrated webcam was used to capture and recognise the participants' emotions, while the external cameras used screen-recording software (Silverback version 2.0) to capture facial expressions of the participants and record the complete session. Raters for validating our software used the recorded video.

3.4 Procedure

Each participant signed the agreement form before his or her session started. Participants individually performed all five tasks in a single session of about 20 min. The session was conducted in a silent room with good lighting condition. The moderator of the session was present in the room but did not intervene. All sessions were conducted on two consecutive days. The participants were requested not to talk to each other in between sessions so that they could not influence each other. The moderator gave a short instruction at the beginning of each task. For example, participants were asked to show mild and not too intense expressions while mimicking the emotions. All tasks were recorded and captured by our software. After the session, each participant filled out an online questionnaire to gather participants' opinions about their learning experience and the setup of the study.

3.5 Validation

Two expert raters analysed the recorded video streams to provide a validation reference for the software output. The raters, both associate professors at the psychology department of the Open University of the Netherlands, were invited to individually rate the emotions of the participants' in the recorded video streams. Both raters are familiar and skilled with using the Facial Action Coding System [20].

Firstly, they received an instruction package for doing individual ratings of participants' emotions in one out of ten video recordings. Secondly, both raters participated in a training session together with the main researcher where ratings of this first participant were discussed to identify possible issues with the rating task and to improve common understanding of the rating categories. Thirdly, raters resumed their individual ratings of participants' emotions in the nine remaining video streams. Fourthly, they participated in a negotiation session together with the main researcher where all ratings were discussed to check whether negotiation about dissimilar ratings could lead to an agreement or to sustained disagreement. Finally, the final ratings resulting from the negotiation session were taken as input for the data analysis. The data of the training session were also included in the final analysis. The raters received: 1) a laptop, 2) a user manual, 3) an instruction guide on how to use ELAN, which is a professional tool for making complex annotations on video and audio resources, and 4) an excel file with ten data sheets; each of which represented the participant's information.

4 Analysis of data and results

In this section, we will first describe how to calculate the total sample size for this study. We then explain the results of the raters. Finally, we explain the agreement between requested emotions and the recognised emotions by the software.

4.1 The required sample size

We used G*Power tool [21], which is a tool to compute statistical power analyses for several statistical tests. We then applied a "t-test" with "correlation analysis of point biserial model" and "a priori" to compute required sample size with given "alpha (significance level)", "power", and "effect size" to realize the total sample size of this study. We used the following input parameters: one tail, effect size = .11, alpha error probability = .05, and the [power (1 - beta error probability)] = .95; so we used beta = 0.05, Type II. The total sample size required for this study appeared to be 885 occurrences with the actual power of .95. We used 1000 occurrences for sampling the 'requested emotions', thus this criterion was met.

4.2 Results of the raters for recognising emotions

Hereafter, we describe how the raters detected participants' emotions from their recorded video streams. The disagreement between the raters, which was 34% before the negotiation session, was reduced to 22% at the end of the negotiation session. In order to determine consistency among raters, we performed the cross-tabulation between the raters and also inter-rater reliability analysis using the Kappa ($\varkappa$) statistic approach [44]. The $\varkappa$ value in statistics can measure inter-rater agreement for qualitative items. We calculated and presented the $\varkappa$ value for the original ratings before negotiation. We have 1000 displayed emotions (see Table 1) rated by two raters as being one of the six basic emotions and the neutral emotion. The cross-tabulation data are given in Table 1. Each recognised emotion by the rater 1 is separated into two rows that intersect with the recognised emotions by the rater 2. The first row indicates the number of occurrences of the recognised emotion and the second row displays the percentage of agreement about the identified emotions. In addition to the '$\varkappa$' value, we also calculated overall agreement ('α') for each table. This 'α' value is the average of each diagonal in the related tables, which is calculated based on the uniform distribution of emotions. For instance the 'α' value in Table 1 is calculated as: $\alpha = (90.6 + 53.3 + 53.3 + 39.7 + 68.2 + 73.4 + 95.2)/7 = 67.7$.

Cross-tabulation analysis between the raters indicates that the neutral expression has the highest agreement (95.2%) and the fear expression has the lowest agreement between them (39.7%) (Table 1). According to Murthy and Jadon [54], people have more difficulty in recognising fear facial expression, which clarifies why the most confused expression is fear. Sadness is the next confused category, which is often recognised as neutral (26.7%). Analysis of the $\varkappa$ statistic underlines the high degree of agreement among the raters. The inter-rater reliability was calculated to be $\varkappa = .715$ ($p < 0.001$), which qualifies as a substantial agreement among raters according to the interpretation of $\varkappa$ values by Landis & Koch [44].

4.3 Emotion recognition by the software

Table 2 shows the requested emotions of participants contrasted with software recognition results. These numbers are taken from all 1000 emotions (10 test persons displaying 100 emotions each) including the cases that one or more of the raters judged that the test person was unable to mimic the requested emotion correctly. Each requested emotion is separated into two rows that intersect with the recognised emotions by the software. Our software has the highest recognition rate for the happy expression (93.3%) and the lowest recognition rate for the fear expression (43.8%) (See Table 2).

Table 1　Rater1 * Rater2 Cross-tabulation – All 1000 emotions are rated by both raters. ($\varkappa$ = .715 and α = 67.7%)

		Rater2							Total
		Happy	Sad	Surprise	Fear	Disgust	Anger	Neutral	
Rater1	Happy	**106**	0	1	1	1	0	8	117
		90.6%	0%	0.9%	0.9%	0.9%	0%	6.7%	100%
	Sad	0	**32**	0	1	3	8	16	60
		0%	**53.3%**	0%	1.7%	5%	13.3%	26.7%	100%
	Surprise	9	0	**57**	8	2	1	30	107
		8.4%	0%	**53.3%**	7.5%	1.9%	0.9%	28%	100%
	Fear	0	0	16	**23**	14	0	5	58
		0%	0%	27.6%	**39.7%**	24.1%	0%	8.6%	100%
	Disgust	0	3	2	2	**58**	8	12	85
		0%	3.5%	2.4%	2.4%	**68.2%**	9.4%	14.1%	100%
	Anger	1	6	1	1	6	**69**	10	94
		1.1%	6.4%	1.1%	1.1%	6.4%	**73.4%**	10.5%	100%
	Neutral	6	4	5	0	1	7	**456**	479
		1.3%	0.8%	1%	0%	0.2%	1.5%	**95.2%**	100%
Total		122	45	82	36	85	93	537	1000

Please note that the obtained differences between software and requested emotions are not necessarily software faults but could also indicate that participants were sometimes unable to mimic the requested emotions (30.6%). The software had in particular problems to distinguish sad from neutral, fear from neutral, anger, disgust, and surprise, disgust from anger and neutral, anger from disgust and sad. Error rates of the software are typically between 0.8% and 31.1%.

The numbers in Table 2 show that all six basic emotions and the neutral one have different distributions for being confused as to the other emotions. In other words, they have different discrimination rates. Apart from neutral, the emotions that are best discriminated from other ones are happiness, surprise, and anger. Happiness has the highest accuracy rate of 93.3% and is not confused with fear and anger at all; surprise has the next highest accuracy rate of 86.3% and is not

Table 2　Requested emotions and emotions recognised by the software – These numbers are taken from all 1000 emotions including 'unable to mimic' by the participants ($\varkappa$ = .716 and α = 71.5%)

		Recognised Emotion by the Software							Total
		Happy	Sad	Surprise	Fear	Disgust	Anger	Neutral	
Requested Emotions	Happy	**112**	1	2	0	2	0	3	120
		93.3%	0.8%	1.7%	0%	1.7%	0%	2.5%	100%
	Sad	2	**43**	1	4	5	7	28	90
		2.2%	**47.8%**	1.1%	4.4%	5.6%	7.8%	31.1%	100%
	Surprise	0	0	**69**	2	5	1	3	80
		0%	0%	**86.3%**	2.5%	6.3%	1.3%	3.8%	100%
	Fear	1	5	8	**35**	8	9	14	80
		1.1%	6.3%	10.0%	**43.8%**	10.0%	11.3%	17.5%	100%
	Disgust	4	1	4	3	**64**	8	6	90
		4.5%	1.0%	4.5%	3.3%	**71.1%**	8.9%	6.7%	100%
	Anger	3	5	4	2	11	**55**	0	80
		3.6%	6.3%	5.0%	2.5%	13.8%	**68.8%**	0%	100%
	Neutral	5	17	10	6	4	5	**413**	460
		1.1%	3.7%	2.2%	1.2%	0.9%	1.1%	**89.8%**	100%
Total		127	72	98	52	99	85	467	1000

confused with happiness and sadness at all. The most difficult emotion is fear, which scores only 43.8% and is easily confused with neutral 17.5%, anger 11.3%, surprise 10.0%, disgust 10.0%, sadness 6.3% and happiness 1.1%, respectively. This is in accordance with Murthy and Jadon [54] and Zhang [83]. Moreover, Murthy and Jadon [54] states that sadness, disgust, and anger are difficult to distinguish from each other and are therefore often wrongly classified.

Taking the raters' analysis results as a reference Table 3 shows that the participants were able to mimic requested emotions correctly in 69.4% of the occurrences. In 200 occurrences (20%) there was disagreement between raters. In the remaining 10.6% of the cases, the raters agreed that participants were unable to mimic requested emotions (106 times). Participants are best at mimicking neutral (87.4%) and worst at mimicking fear correctly (21.3%), which is in accordance with Murthy and Jadon [54].

Table 4 shows the requested emotions of participants contrasted with software recognition results while excluding both the 'unable to mimic' records and the records on which the raters disagreed with the dataset. We, therefore, re-calculated the results of each emotion separately and in total.

In 306 out of 1000 cases at least one of the raters indicated that the participants were 'unable to mimic' the requested emotions properly. We only summed occurrences when both raters agreed that the displayed emotion was the same as the requested emotion'. The results for all emotions move toward positive changes. Table 5 shows the comparison between the accuracy of results of Tables 2 and 4.

The overall accuracy of 83.2% and the associated **x** value of .837 are the final results that fully rely on the comparison of requested emotions and recognised emotions.

4.4 Comparison of our software output with the extended Cohn-Kanade database

Table 6 shows the labelled emotions of the Cohn-Kanade subjects contrasted with the FURIA classifier algorithm of our software. These numbers are taken from all 432 labelled emotions of the subjects including the 432 rotated images of the subjects (in total = 864 emotions). Each labelled emotion is separated into two rows that intersect with the recognised emotions through the FURIA classifier algorithm. The FURIA classifier algorithm in our software has the highest recognition rate for the surprise expression (95.2%) and the lowest recognition rate for the fear expression (34%).

Based on our calculation using WEKA, the correctly classified instances of the Cohn-Kanade database are 678 instances with the accuracy rate of 78.5% and the incorrectly classified instances are 186 instances with the accuracy 21.5%. The Kappa value is therefore **x** = .737. Table 7 shows the comparison between the accuracy results of Tables 4 and 6 for each independent emotion.

Table 3 Raters' agreements and disagreements about 1000 mimicked emotions

	Happy	Sad	Surprise	Fear	Disgust	Anger	Neutral	Total
Raters agree:	102	24	50	17	47	52	402	694
Able to mimic	85%	26.7%	62.5%	21.3%	52.2%	65%	87.4%	69.4%
Raters disagree:	16	31	22	24	34	22	51	200
Able/unable to mimic	13.3%	34.4%	27.5%	30%	37.8%	27.5%	11.1%	20%
Raters agree: Unable	2	35	8	39	9	6	7	106
to mimic	1.7%	38.9%	10%	48.8%	10%	7.5%	1.5%	10.6%
								100%

Table 4 Requested emotions and recognised emotions by the software – These numbers are taken by the raters from 694 emotions of the participants that were able to mimic the requested emotions ($\varkappa = .837$ and $\alpha = 83.2\%$)

		Recognised Emotion by the Software							Total
		Happy	Sad	Surprise	Fear	Disgust	Anger	Neutral	
Requested Emotions	Happy	**99**	1	1	0	1	0	0	102
		97.0%	1.0%	1.0%	0.0%	1.0%	0.0%	0.0%	100%
	Sad	1	**18**	0	2	0	0	3	24
		4.2%	**75.0%**	0.0%	8.3%	0.0%	0.0%	12.5%	100%
	Surprise	0	0	**47**	1	1	1	0	50
		0.0%	0.0%	**94.0%**	2.0%	2.0%	2.0%	0.0%	100%
	Fear	0	0	2	**12**	0	0	3	17
		0.0%	0.0%	11.8%	**70.6%**	0.0%	0.0%	17.6%	100%
	Disgust	0	0	0	0	**39**	6	2	47
		0.0%	0.0%	0.0%	0.0%	**83.0%**	12.8%	4.2%	100%
	Anger	3	3	3	1	5	**37**	0	52
		5.8%	5.8%	5.8%	1.9%	9.5%	**71.2%**	0%	100%
	Neutral	4	8	7	5	4	5	**369**	402
		1.0%	2.0%	1.8%	1.2%	1.0%	1.2%	**91.8%**	100%
Total		107	30	60	21	50	49	377	694

The results show that the accuracy of our software overcomes the accuracy of the Cohn-Kanade database. While, the precisions of sad, fear, and anger emotions show significant increases, the precisions of happy, disgust and neutral emotions show the small improvements. The surprise emotion shows less precision in our results. This might be the case that the total sample size required for this study appeared to be minimum 885 occurrences with the actual power of .95. However, for analysing the Cohn-Kanade database, we used only 432 frontal faces as well as 432 rotated faces, thus this criterion was not met properly.

5 Discussion

This study presented an analysis for establishing the accuracy of facial emotion recognition based on a fuzzy logic model. The result showed that $\varkappa = .837$ and an average accuracy $\alpha = 83.2\%$ based on the comparison of recognised emotions and requested emotions. The data show that most intensive emotions (e.g., happiness, surprise) can be detected better than the less intensive emotions except neutral and fear. This is in accordance with Murthy and Jadon

Table 5 The comparison between the accuracy results of Tables 2 and 4. Each emotion is independently compared

	All 1000 emotions	Only 694 able to mimic emotions
Happy	93.3%	97.1%
Sad	47.8%	75.0%
Surprise	83.6%	94.0%
Fear	43.8%	70.6%
Disgust	71.1%	83.0%
Anger	68.8%	71.2%
Neutral	89.8%	91.8%
Average accuracy α	71.5%	83.2%

Table 6 Recognition rate of the FURIA classifier algorithm our software over the labelled emotions of the Cohn-Kanade database – These numbers are taken from all 432 labelled emotions of the subjects including the 432 rotated images of the subjects

		Recognised Emotion by the Software							Total
		Happy	Sad	Surprise	Fear	Disgust	Anger	Neutral	
Requested Emotions	Happy	**131**	0	0	0	3	0	4	138
		94.9%	0.0%	0.0%	0.0%	2.2%	%	2.9%	100%
	Sad	0	**20**	0	0	0	7	27	54
		0.0%	**37.0%**	0.0%	0.0%	0.0%	13.0%	50.0%	100%
	Surprise	1	0	**160**	3	1	0	3	168
		0.6%	0.0%	**95.2%**	1.8%	0.6%	0.0%	1.8%	100%
	Fear	3	2	3	**17**	3	0	22	50
		6%	4%	6%	**34.0%**	6.0%	0.0%	44.0%	100%
	Disgust	7	0	0	0	**90**	7	14	118
		6.0%	0.0%	0.0%	0.0%	**76.2%**	6.0%	11.8%	100%
	Anger	3	10	0	1	13	**52**	11	90
		3.3%	11.1%	0.0%	1.1%	14.4%	**57.9%**	12.2%	100%
	Neutral	6	8	1	2	9	12	**208**	246
		2.4%	3.2%	0.4%	0.8%	3.7%	4.9%	**84.6%**	100%
Total		151	40	164	23	119	78	289	864

[54] and Zhang [83], who found that the most difficult emotion to mimic accurately is fear. Moreover, this result expresses that fear is differently interpreted from other basic facial emotions. Furthermore, our data analysis confirms Murthy's [54] finding that sadness, disgust, and anger are difficult to distinguish from each other and are therefore often wrongly classified. Anger and disgust share many similar facial actions [20] and that is probably the reason why they are often confused. In 137 cases of disgust from joint Tables 2 and 4, 14 cases are detected as anger. In 132 cases of anger from Tables 2 and 4, 16 cases are detected as disgust. Hence confusion of anger and disgust is well over 8.9%.

Some potential limitations of the study should be pointed out. First, we have considered only six basic emotions and the neutral emotion in this study, although a larger diversity might be opportune. Nevertheless, the fuzzy logic approach could be easily extended to more emotions provided that an annotated reference database is available. Second, to validate the fuzzy-logic approach we used the recorded data of non-actors. A previous study by Krahmer and Swerts has shown that actors, although they evidently have better acting skills than laymen, will produce more realistic (i.e., authentic, spontaneous)

Table 7 The comparison between the recognition results of Tables 4 and 6. Each emotion is independently compared

	Results of Table 6: Cohn-Kanade database based on the FURIA classifier	Results of Table 4: Our database based on the FURIA classifier
Happy	94.9%	97.1%
Sad	37.0%	75.0%
Surprise	95.2%	94.0%
Fear	34.0%	70.6%
Disgust	76.2%	83.0%
Anger	57.9%	71.2%
Neutral	84.6%	91.8%

expressions [41]. Third, given our sample of medium-aged participants, we did not take into account participants' age as a disturbing factor. Existing research shows that youngsters and older adults are not equally good at mimicking different basic emotions, e.g., older adults are less good at mimicking sadness and happiness than youngsters, but older adults can mimic disgust in a better way than youngsters do [30]. Likewise, potential gender differences have not been taken into account.

6 Conclusion

The presented approach to fuzzy-logic based emotion recognition offers high quality, reliable recognition, and categorisation of emotions. The approach fulfils the requirements of being 1) an unobtrusive approach, with 2) an objective method that can be verified by researchers, 3) which requires inexpensive and ubiquitous equipment (webcam), and 4) which outperforms existing approaches. Compared to our previous study with the accuracy of 72% and less reliability [7], this study achieve a 83.2% average accuracy (α) level, which is comparable with human performance [15, 55]. Moreover, multiple faces in a picture can detect at the same time. Furthermore, being compliant with the RAGE software architecture, the emotion recognition component created in this study can be easily ported to a variety of game engines and e-learning environments.

Emotion recognition technology can now be easily added to educational games or e-learning environments to enhance overall support for the learning. It opens up new possibilities to including the learners' emotional states in the user profiles needed for adaptive and personalised feedback, and to the dedicated training of communication skills and other soft skills are heavily rely on emotion [28]. This technology can also be easily used in other domains, such as the healthcare. For instance, this technology can be used in social adaptation skills for children with autism spectrum disorder [33].

Appendix 1

FURIA Fuzzy Rules (37 Rules):

- **Rule 1:** (Cosine30 in [159.608, 160.424, inf, inf]) and (Cosine35 in [30.0655, 30.2536, inf, inf]) = > Emotions = Happy (CF = 0.97)
- **Rule 2:** (Cosine31 in [−inf, −inf, 9.85016, 9.96286]) and (Cosine51 in [35.5078, 35.5856, inf, inf]) and (Cosine10 in [22.9748, 23.0757, inf, inf]) and (Cosine0 in [−inf, −inf, 24.8358, 24.9599]) = > Emotions = Happy (CF = 0.98)
- **Rule 3:** (Cosine32 in [−inf, −inf, 13.8022, 14.6793]) and (Cosine34 in [25.1647, 25.4542, inf, inf]) and (Cosine50 in [3.7724, 3.79234, inf, inf]) and (Cosine48 in [64.551, 66.8476, inf, inf]) = > Emotions = Happy (CF = 0.9)

- **Rule 4:** (Cosine31 in [−inf, −inf, 12.4074, 16.2636]) and (Cosine35 in [25.6426, 27.0878, inf, inf]) and (Cosine37 in [21.4168, 21.5797, inf, inf]) and (Cosine0 in [−inf, −inf, 21.1748, 21.9187]) => Emotions = Happy (CF = 0.95)
- **Rule 5:** (Cosine32 in [−inf, −inf, 5.15255, 5.16574]) and (Cosine32 in [4.97626, 5.05265, inf, inf]) => Emotions = Happy (CF = 0.66)
- **Rule 6:** (Cosine34 in [−inf, −inf, 15.1591, 15.3646]) and (Cosine42 in [−inf, −inf, 73.6283, 74.5163]) and (Cosine32 in [17.9646, 17.9917, inf, inf]) and (Cosine13 in [68.61, 70.3929, inf, inf]) and (Cosine37 in [−inf, −inf, 22.5483, 22.8643]) => Emotions = Sad (CF = 0.94)
- **Rule 7:** (Cosine33 in [142.341, 142.606, inf, inf]) and (Cosine44 in [51.9081, 52.9607, inf, inf]) and (Cosine50 in [−inf, −inf, 3.15842, 3.18719]) and (Cosine31 in [14.1431, 14.4425, inf, inf]) => Emotions = Sad (CF = 0.82)
- **Rule 8:** (Cosine46 in [−inf, −inf, 84.9609, 85.074]) and (Cosine51 in [36.248, 36.6846, inf, inf]) => Emotions = Sad (CF = 0.81)
- **Rule 9:** (Cosine42 in [−inf, −inf, 54.362, 54.4623]) and (Cosine3 in [49.4672, 53.9726, inf, inf]) and (Cosine18 in [48.7314, 48.7806, inf, inf]) => Emotions = Sad (CF = 0.84)
- **Rule 10:** (Cosine1 in [−inf, −inf, 6.82602, 7.03498]) and (Cosine15 in [14.5889, 15.0512, inf, inf]) => Emotions = Sad (CF = 0.53)
- **Rule 11:** (Cosine30 in [−inf, −inf, 126.304, 149.554]) and (Cosine34 in [33.9611, 34.0477, inf, inf]) => Emotions = Surprise (CF = 0.99)
- **Rule 12:** (Cosine8 in [40.2461, 40.3331, inf, inf]) and (Cosine12 in [−inf, −inf, 45.1865, 45.2721]) => Emotions = Surprise (CF = 0.98)
- **Rule 13:** (Cosine17 in [17.0409, 17.0477, inf, inf]) and (Cosine17 in [−inf, −inf, 17.1027, 17.1054]) => Emotions = Surprise (CF = 0.68)
- **Rule 14:** (Cosine48 in [−inf, −inf, 42.3974, 43.2472]) => Emotions = Surprise (CF = 0.6)
- **Rule 15:** (Cosine26 in [38.6598, 40.2364, inf, inf]) and (Cosine50 in [3.61388, 3.62035, inf, inf]) and (Cosine19 in [86.8202, 87.1376, inf, inf]) and (Cosine32 in [−inf, −inf, 22.3228, 22.4199]) and (Cosine44 in [48.7314, 48.8141, inf, inf]) => Emotions = Fear (CF = 0.9)
- **Rule 16:** (Cosine49 in [−inf, −inf, 95.8395, 96.2791]) and (Cosine49 in [93.3419, 94.0108, inf, inf]) and (Cosine8 in [35.0195, 35.6748, inf, inf]) and (Cosine30 in [112.548, 127.911, inf, inf]) => Emotions = Fear (CF = 0.84)
- **Rule 17:** (Cosine33 in [113.647, 120.77, inf, inf]) and (Cosine35 in [19.8265, 20.2393, inf, inf]) and (Cosine22 in [41.6335, 41.8665, inf, inf]) and (Cosine51 in [34.8844, 35.0913, inf, inf]) and (Cosine12 in [−inf, −inf, 56.658, 56.7193]) => Emotions = Fear (CF = 0.87)
- **Rule 18:** (Cosine26 in [−inf, −inf, 26.9489, 27.5973]) and (Cosine8 in [−inf, −inf, 25.3267, 26.1046]) and (Cosine21 in [59.7436, 60.0864, inf, inf]) => Emotions = Disgust (CF = 0.96)
- **Rule 19:** (Cosine26 in [−inf, −inf, 35.8377, 36.2538]) and (Cosine30 in [−inf, −inf, 151.318, 151.76]) and (Cosine35 in [18.6981, 18.7149, inf, inf]) and (Cosine8 in [−inf, −inf, 25.9913, 26.0754]) and (Cosine50 in [3.1012, 3.21448, inf, inf]) and (Cosine51 in [−inf, −inf, 40.4826, 40.6054]) => Emotions = Disgust (CF = 0.96)
- **Rule 20:** (Cosine26 in [−inf, −inf, 35.4509, 35.8195]) and (Cosine30 in [−inf, −inf, 160.168, 160.675]) and (Cosine42 in [91.4652, 91.8476, inf, inf]) and (Cosine9 in [28.8713, 31.4661, inf, inf]) and (Cosine16 in [−inf, −inf, 151.557, 151.794]) => Emotions = Disgust (CF = 0.94)

- **Rule 21:** (Cosine28 in [−inf, −inf, 38.1572, 38.3107]) and (Cosine31 in [7.43669, 10.1433, inf, inf]) and (Cosine26 in [−inf, −inf, 30.9638, 31.4521]) and (Cosine4 in [−inf, −inf, 46.4321, 46.5831]) and (Cosine45 in [74.4759, 75.0686, inf, inf]) => Emotions = Disgust (CF = 0.94)
- **Rule 22:** (Cosine28 in [−inf, −inf, 38.1572, 38.3782]) and (Cosine32 in [8.70181, 8.71473, inf, inf]) and (Cosine9 in [35.9909, 36.1861, inf, inf]) and (Cosine8 in [−inf, −inf, 27.817, 27.929]) and (Cosine14 in [−inf, −inf, 46.8177, 47.5595]) and (Cosine8 in [17.2892, 18.9069, inf, inf]) => Emotions = Disgust (CF = 0.93)
- **Rule 23:** (Cosine28 in [−inf, −inf, 38.1572, 39.6233]) and (Cosine32 in [15.3773, 15.6442, inf, inf]) and (Cosine23 in [−inf, −inf, 90, 92.4002]) and (Cosine6 in [34.6599, 35.2154, inf, inf]) and (Cosine2 in [144.482, 144.492, inf, inf]) => Emotions = Disgust (CF = 0.9)
- **Rule 24:** (Cosine10 in [−inf, −inf, 26.4547, 27.2133]) and (Cosine17 in [13.8931, 14.1754, inf, inf]) and (Cosine11 in [−inf, −inf, 117.723, 118.822]) and (Cosine7 in [112.807, 113.395, inf, inf]) and (Cosine0 in [−inf, −inf, 24.2531, 28.1416]) => Emotions = Disgust (CF = 0.86)
- **Rule 25:** (Cosine7 in [115.258, 117.446, inf, inf]) and (Cosine34 in [−inf, −inf, 17.2653, 19.2307]) and (Cosine30 in [152.184, 152.257, inf, inf]) and (Cosine36 in [141.526, 141.77, inf, inf]) => Emotions = Anger (CF = 0.93)
- **Rule 26:** (Cosine11 in [122.859, 124.442, inf, inf]) and (Cosine22 in [33.0558, 33.6901, inf, inf]) and (Cosine8 in [−inf, −inf, 21.139, 21.5363]) => Emotions = Anger (CF = 0.9)
- **Rule 27:** (Cosine11 in [114.76, 115.918, inf, inf]) and (Cosine34 in [−inf, −inf, 17.0449, 17.049]) and (Cosine39 in [−inf, −inf, 137.883, 138.125]) and (Cosine19 in [86.0353, 86.1859, inf, inf]) => Emotions = Anger (CF = 0.82)
- **Rule 28:** (Cosine11 in [123.048, 124.509, inf, inf]) and (Cosine51 in [−inf, −inf, 33.6707, 33.6995]) and (Cosine18 in [52.6819, 53.231, inf, inf]) => Emotions = Anger (CF = 0.83)
- **Rule 29:** (Cosine34 in [−inf, −inf, 22.8636, 23.1986]) and (Cosine31 in [−inf, −inf, 10.1964, 10.3628]) and (Cosine46 in [−inf, −inf, 107.21, 107.447]) and (Cosine47 in [−inf, −inf, 3.29176, 3.4244]) => Emotions = Anger (CF = 0.82)
- **Rule 30:** (Cosine5 in [−inf, −inf, 68.4986, 68.8774]) and (Cosine35 in [−inf, −inf, 18.6981, 18.7149]) and (Cosine13 in [−inf, −inf, 66.8664, 67.7345]) and (Cosine51 in [−inf, −inf, 39.3543, 39.4115]) => Emotions = Anger (CF = 0.92)
- **Rule 31:** (Cosine34 in [−inf, −inf, 26.5651, 26.7055]) and (Cosine20 in [39.2894, 39.6233, inf, inf]) and (Cosine45 in [−inf, −inf, 78.9618, 78.9745]) and (Cosine38 in [19.9831, 20.0674, inf, inf]) and (Cosine8 in [−inf, −inf, 35.4327, 36.92]) and (Cosine30 in [−inf, −inf, 156.633, 157.079]) and (Cosine44 in [38.5527, 38.6598, inf, inf]) => Emotions = Neutral (CF = 0.98)
- **Rule 32:** (Cosine35 in [−inf, −inf, 27.2363, 27.3464]) and (Cosine28 in [38.1572, 38.6598, inf, inf]) and (Cosine6 in [38.9806, 39.1219, inf, inf]) and (Cosine50 in [−inf, −inf, 3.61388, 3.64164]) and (Cosine45 in [−inf, −inf, 79.5199, 79.6952]) => Emotions = Neutral (CF = 0.95)
- **Rule 33:** (Cosine34 in [−inf, −inf, 28.1313, 28.9551]) and (Cosine28 in [38.1572, 38.4537, inf, inf]) and (Cosine9 in [39.7376, 39.8419, inf, inf]) and (Cosine19 in [−inf, −inf, 86.6335, 87.0364]) and (Cosine50 in [−inf, −inf, 3.61207, 3.62575]) => Emotions = Neutral (CF = 0.93)
- **Rule 34:** (Cosine34 in [−inf, −inf, 26.8913, 26.9486]) and (Cosine20 in [33.6901, 35.8195, inf, inf]) and (Cosine12 in [−inf, −inf, 58.3317, 58.6666]) and (Cosine22 in

[−inf, −inf, 38.7933, 39.0939]) and (Cosine33 in [−inf, −inf, 152.969, 152.971]) and (Cosine43 in [50.7021, 50.8786, inf, inf]) => Emotions = Neutral (CF = 0.94)

- **Rule 35:** (Cosine33 in [88.1516, 121.909, inf, inf]) and (Cosine28 in [38.1572, 38.6238, inf, inf]) and (Cosine51 in [−inf, −inf, 35.422, 36.0181]) and (Cosine43 in [−inf, −inf, 51.481, 51.6839]) and (Cosine32 in [13.4486, 13.6831, inf, inf]) and (Cosine26 in [−inf, −inf, 43.2643, 43.3844]) and (Cosine36 in [−inf, −inf, 146.691, 146.723]) => Emotions = Neutral (CF = 0.97)

- **Rule 36:** (Cosine35 in [−inf, −inf, 29.0332, 32.685]) and (Cosine8 in [26.5308, 26.5651, inf, inf]) and (Cosine29 in [−inf, −inf, 85.6013, 85.9144]) and (Cosine44 in [−inf, −inf, 51.953, 52.0023]) and (Cosine5 in [−inf, −inf, 75.7986, 76.0875]) and (Cosine18 in [−inf, −inf, 61.0736, 64.9831]) => Emotions = Neutral (CF = 0.95)

- **Rule 37:** (Cosine33 in [118.764, 120.426, inf, inf]) and (Cosine27 in [−inf, −inf, 57.0948, 64.7327]) and (Cosine49 in [104.517, 104.534, inf, inf]) and (Cosine51 in [−inf, −inf, 33.918, 34.2595]) and (Cosine5 in [73.3444, 73.7261, inf, inf]) and (Cosine33 in [−inf, −inf, 136.4, 149.673]) => Emotions = Neutral (CF = 0.96)

References

1. Ali H, Hariharan M, Yaacob S, Adom AH (2015) Facial Emotion Recognition Using Empirical Mode Decomposition. Expert Syst Appl 42(3):1261–1277
2. Anisetti M, Bellandi V, Damiani E, Jeon G, Jeong J, Sellitto S (2009) Emotional state inference using face related features. International Conference on Interfaces and Human Computer Interaction, Porto
3. Arroyo I, Woolf B, Cooper D, Burleson W, Muldner K, Christopherson R (2009) Emotion Sensors Go To School. Artificial Intelligence in Education 1(1):18–37
4. Bachiller C, Hernandez C, Sastre J (2010) Collaborative Learning, Research and Science Promotion in a Multidisciplinary Scenario: Information and Communications Technology and Music. Proceedings of the International Conference on Engineering Education, Gliwice, pp 1–8
5. Bahreini K, Elci A (2008a) A New Software Architecture for J2EE Enterprise Environments via Semantic Access to Web Sources for Web Mining by Distributed Intelligent Software Agents. Proceedings of the 32nd Annual IEEE International Computer Software and Applications Conference (COMPSAC), Turku, pp 902–907. https://doi.org/10.1109/COMPSAC
6. Bahreini K, Elci A (2008b) SDISSASA: A Multiagent-Based Web Mining via Semantic Access to Web Resources in Enterprise Architecture. Proceedings of the 32nd Annual {IEEE} International Computer Software and Applications Conference (COMPSAC), Turku, pp 553–558. https://doi.org/10.1109/COMPSAC
7. Bahreini K, Nadolski RJ, Westera W (2016a) Towards Multimodal Emotion Recognition in E-Learning Environments. Interact Learn Environ 24(3):590–605
8. Bahreini K, Nadolski R, Westera W (2016b) Towards Real-Time Speech Emotion Recognition for Affective E-Learning. Educ Inf Technol 21(5):1367–1386. https://doi.org/10.1007/s10639-015-9388-2
9. Bahreini K, Nadolski R, Westera W (2016c) Data Fusion for Real-time Multimodal Emotion Recognition through Webcams and Microphones in E-Learning. International Journal of Human-Computer Interaction 32(5):415–430. Taylor & Francis. https://doi.org/10.1080/10447318.2016.1159799

10. Bahreini K, Nadolski R, Westera W (2017) Communication Skills Training Exploiting Multimodal Emotion Recognition. Interact Learn Environ 25(8):1065–1082. Routledge. https://doi.org/10.1080/10494820.2016.1247286
11. Ben Ammar M, Neji M, Alimi AM, Gouardères G (2010) The Affective Tutoring System. Expert Syst Appl 37(4):3013–3023
12. Bettadapura V (2012) Face expression recognition and analysis: the state of the art. Journal of CoRR, abs/1203.6722
13. Brusilovsky P (1994a) Student Model-Centered Architecture for Intelligent Learning Environments. Proceedings of Fourth International Conference on User Modeling, Hyannis, pp 31–36
14. Brusilovsky P (1994b) The Construction and Application of Student Models in Intelligent Tutoring Systems. J Comput Syst Sci Int 32(1):70–89
15. Burkhardt F, Paeschke A, Rolfes M, Sendlmeier W, Weiss B (2005) A Database of German Emotional Speech. Proceedings of the Inter Speech, Lissabon, pp 1517–1520
16. Chung-Hsien W, Ze-Jing C, Yu-Chung L (2006) Emotion Recognition from Text Using Semantic Labels and Separable Mixture Models. Journal of ACM Transactions on Asian Language Information Processing (TALIP) 5(2):165–183. ACM, New York. https://doi.org/10.1145/1165255.1165259
17. D'Mello SK, Graesser AC (2012) AutoTutor and Affective AutoTutor: Learning by Talking with Cognitively and Emotionally Intelligent Computers that Talk Back. ACM Transactions on Interactive Intelligent Systems 2(4):1–39
18. De Gloria A, Bellotti F, Berta R (2014) Serious Games for Education and Training. International Journal of Serious Games 1(1). Retrieved from http://journal.seriousgamessociety.org/. Accessed 1 Feb 2019
19. Devi JS, Srinivas Y, Nandyala SP (2014) Automatic Speech Emotion and Speaker Recognition based on Hybrid GMM and FFBNN. International Journal on Computational Sciences & Applications (IJCSA) 4(1):35–42
20. Ekman P, Friesen WV (1978) Facial action coding system: investigator's guide. Consulting Psychologists Press
21. Erdfelder E, Faul F, Buchner A (1996) GPOWER: A general power analysis program. Behav Res Methods Instrum Comput 28:1–11
22. Esau N, Wetzel E, Kleinjohann L, Kleinjohann B (2007) Real-time facial expression recognition using a fuzzy emotion model. In IEEE International Fuzzy Systems Conference, FUZZ-IEEE 2007
23. Faundez-Zanuy M, Espinosa-Duró V, Ortega J (2005) A low-cost Webcam & personal computer opens doors. IEEE Aerosp Electron Syst Mag 20:23–26
24. Feidakis M, Daradoumis T, Caballe S (2011) Emotion measurement in intelligent tutoring systems: what, when and how to measure. Third International Conference on Intelligent Networking and Collaborative Systems, 807–812
25. Goetz T, Lüdtke O, Ulrike EN, Keller MM, Lipnevich AA (2013) Characteristics of Teaching and Students' Emotions in the Classroom: Investigating Differences Across Domains. Contemp Educ Psychol 38(4):383–394. https://doi.org/10.1016/j.cedpsych.2013.08.001
26. Gu WF, Venkatesh YV, Xiang C (2010) A Novel Application of Self-Organizing Network for Facial Expression Recognition from Radial Encoded Contours. Soft Comput 14:113–122
27. Gunes H, Pantic M (2010) Automatic, Dimensional and Continuous Emotion Recognition. International Journal of Synthetic Emotions 1(1):68–99
28. Hager PJ, Hager P, Halliday J (2006) Recovering informal learning: wisdom, judgment and community. Springer
29. Huhn J, Hullermeier E (2009a) FURIA: An Algorithm for Unordered Fuzzy Rule Induction. Data Min Knowl Disc 19(3):293–319
30. Huhnel I, Fölster M, Werheid K, Hess U (2014) Empathic reactions of younger and older adults: No age-related decline in affective responding. J Exp Soc Psychol 50:136–143
31. Hussein MS, Hussain MS, AlZoubi O, Calvo RA, D'Mello SK (2011). Affect detection from multichannel physiology during learning sessions with AutoTutor. Artificial Intelligence in Education, 6738, 131-138. Auckland: Springer, LNAI
32. Ilbeygi M, Hoseini HS (2012) A Novel Fuzzy Facial Expression Recognition System Based on Facial Feature Extraction from Color Face Images. Eng Appl Artif Intell 25:130–146
33. Ip HHS, Wong SWL, Chan DFY, Byrne J, Li C, Yuan VSN, Lau KSY, Wong JYW (2018) Enhance Emotional and Social Adaptation Skills for Children with Autism Spectrum Disorder: A Virtual Reality Enabled Approach. Comput Educ 117:1–15. https://doi.org/10.1016/j.compedu.2017.09.010
34. Jaimes A, Sebe N (2007) Multimodal Human–Computer Interaction: A Survey. Computer Vision and Image Understanding. Special Issue on Vision for. Human-Computer Interaction 108(1-2):116–134
35. Jianhua T, Tieniu T, Rosalind WP (2005) Affective computing: a review, affective computing and intelligent interaction. Springer, Berlin, 3784, 981-995

36. Kazmi SB, Qurat-ul-Ain SB, Jaffar MA (2012) Wavelet-Based Facial Expression Recognition Using a Bank of Support Vector Machines. Soft Comput 16(3):369–379

37. Kharat GU, Dudul SV (2009) Emotion recognition from facial expression using neural networks. In: Hippe, ZS, Kulikowski JL (Eds.), Human–Computer Systems Interaction, AISC 60, 207-219

38. Kim DJ (2016) Facial Expression Recognition using ASM-Based Post-Processing Technique. Pattern Recognition and Image Analysis 26(3):576–581. Pleiades Publishing. https://doi.org/10.1134/S105466181603010X

39. King DE (2009) Dlib-ml: A Machine Learning Toolkit. J Mach Learn Res 10:1755–1758

40. Kipp M, Martin JC (2009) Gesture and Emotion: Can Basic Gestural form Features Discriminate Emotions? 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (ACII 2009), 1-8. https://doi.org/10.1109/ACII.2009.5349544

41. Krahmer E, Swerts M (2011) Audiovisual expression of emotions in communication. Philips Research Book Series. Springer Netherlands, 12, 85-106

42. Kushki A, Fairley J, Merja S, King G, Chau T (Oct 2011) Comparison of Blood Volume Pulse and Skin Conductance Responses to Mental and Affective Stimuli at Different Anatomical Sites. Physiol Meas 32(10):1529–1539. https://doi.org/10.1088/0967-3334/32/10/002

43. Kyriakos S, Ekaterini S, Nikos P (2014) A Dynamic Gesture and Posture Recognition System. J Intell Robot Syst 76(2):283–296. Springer Netherlands. https://doi.org/10.1007/s10846-013-9983-7

44. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174

45. Lang G, van der Molen HT (2008) Psychologische gespreksvoering book. Open University of the Netherlands. Heerlen, The Netherlands

46. Lee H, Choi YS, Lee S, Park IP (2012) Towards unobtrusive emotion recognition for affective social communication. Proceedings of the 9th IEEE Consumer Communications and Networking Conference, 260-264

47. Li SZ, Jain A (2011) Handbook of face recognition. The Second Edition. https://doi.org/10.1007/978-0-85729-932-1

48. Li H, Ren F (2009) The Study on Text Emotional Orientation Based on A Three-Dimensional Emotion Space Model. IEEE International Conference on Natural Language Processing and Knowledge Engineering (IEEE NLP-KE), 1-6. Dalian. https://doi.org/10.1109/NLPKE.2009.5313815

49. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I (2010) The Extended Cohn-Kande Dataset (CK+): A Complete Facial Expression Dataset for Action Unit and Emotion-Specified Expression. Paper Presented at the Third IEEE Workshop on CVPR for Human Communicative Behavior Analysis

50. Luo Y, Wu C-M, Zhang Y (2013) Facial Expression Recognition Based on Fusion Feature of PCA and LBP with SVM. International Journal for Light and Electron Optics 124(17):2767–2770

51. Ma C, Prendinger H, Ishizuka M (2005) A chat system based on emotion estimation from text and embodied conversational messengers. In: Proceedings of the International Conference on Entertainment Computing, 535-538

52. Magdin M, Turcani M, Hudec L (2016) Evaluating the Emotional State of a User Using a Webcam. Special Issue on Artificial Intelligence Underpinning 4(1):61–68

53. Marston WM (1917) Systolic Blood Pressure Symptoms of Deception. J Exp Psychol 2(2):117–163

54. Murthy GRS, Jadon RS (2009) Effectiveness of Eigenspaces for facial expression recognition. International Journal of Computer Theory and Engineering 1(5):638–642

55. Nwe T, Foo S, De Silva L (2003) Speech Emotion Recognition Using Hidden Markov Models. Speech Comm 41:603–623

56. Owusu E, Zhan Y, Mao QR (2014) A Neural-Adaboost Based Facial Expression Recognition System. Expert Syst Appl 41:3383–3390

57. Pantic M, Sebe N, Cohn JF, Huang T (2005) Affective Multimodal Human-Computer Interaction. Proceedings of the 13th Annual ACM International Conference on Multimedia, 5, 669-676, Hilton

58. Pekrun R (1992) The impact of emotions on learning and achievement: towards a theory of cognitive/motivational mediators. J Appl Psychol 41:359–376

59. Piana S, Stagliano A, Odone F, Verri A, Camurri A (2014) Real-time Automatic Emotion Recognition from Body Gestures. CoRR, abs/1402.5047. Available online at arxiv.org/abs/1402.5047. Accessed 1 Feb 2019

60. Preeti K (2013) Multimodal Emotion Recognition for Enhancing Human-Computer Interaction. PhD. Dissertation, March 2013, University of Narsee Monjee, Institute of Management Studies, Department of Computer Engineering, Mumbai, India. Available Online at: shodhganga.inflibnet.ac.in/handle/10603/7529

61. Rita F, Ignazio P, Genoveffa T (2012) Wiimote and Kinect: gestural user interfaces add a natural third dimension to HCI. In: Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI'12), 116-123. ACM, New York. https://doi.org/10.1145/2254556.2254580

62. Russell JA (1980) A Circumplex Model of Affect. J Pers Soc Psychol 39:1161–1178

63. Sarrafzadeh A, Alexander S, Dadgostar F, Fan C, Bigdeli A (2008) How do you know that I don't understand? A look at the future of intelligent tutoring systems. Comput Hum Behav 24(4):1342–1363
64. Sebe N (2009) Multimodal Interfaces: Challenges and Perspectives. Journal of Ambient Intelligence and Smart Environments 1(1):23–30
65. Sebe N, Cohen I, Gevers T, Huang TS (2005) Multimodal approaches for emotion recognition: a survey. Proceeding SPIE, Conference Volume 5670, Internet Imaging VI, 56, San Jose. https://doi.org/10.1117/12.600746
66. Shaheen, S, El-Hajj W, Hajj H, Elbassuoni S (2014) Emotion recognition from text based on automatically generated rules. IEEE International Conference on Data Mining. Workshop (ICDMW), 383-392. https://doi.org/10.1109/ICDMW.2014.80
67. Shen L, Wang M, Shen R (2009) Affective E-Learning: Using Emotional Data to Improve Learning in Pervasive Learning Environment. Educational Technology and Society 12(2):176–189
68. Shotton J, Fitzgibbon A, Cook M, Sharp T, Finocchio M, Moore R, … Blake A (2011) Real-time human pose recognition in parts from a single depth image. IEEE Expert procedings (CVPR). Available Online at: research.microsoft.com/apps/pubs/default.aspx?id=145347. Accessed 1 Feb 2019
69. Shyamal P, Hyung P, Paolo B, Leighton C, Mary R (2012) A Review of Wearable Sensors and Systems with Application in Rehabilitation. Journal of Neuro Engineering and Rehabilitation 9(1). https://doi.org/10.1186/1743-0003-9-21
70. Smith TC, Frank E (2016) Statistical Genomics: Methods and Protocols. Chapter Introducing Machine Learning Concepts with WEKA, 353-378. Springer, New York
71. Tettegah, S. Y., Gartmeier, M. (2015). Emotions, Technology, Design, and Learning. First edition. Academic Press. ISBN= 0128018569 and 9780128018569
72. Van der Molen HT, Gramsbergen-Hoogland YH (2005) Communication in Organizations: Basic Skills and Conversation Models. Psychology Press, New York
73. van der Vegt W, Nyamsuren E, Westera W (2016b) RAGE Reusable Game Software Components and Their Integration into Serious Game Engines. In: Kapitsaki GM, de Almeida ES (eds) Bridging with Social-Awareness, 15th International Conference, ICSR 2016, vol 9679. Proceedings, Lecture Notes in Computer Science, Limassol, pp 165–180
74. van der Vegt W, Westera W, Nyamsuren E, Georgiev A, Martínez Ortiz I (2016a) RAGE Architecture for Reusable Serious Gaming Technology Components. International Journal of Computer Games Technology 2016:5680526. https://doi.org/10.1155/2016/5680526
75. Vayrynen E (2014) Emotion recognition from speech using prosodic features. University Of Oulu, Graduate School, Faculty of Information Technology And Electrical Engineering, Department Of Computer Science and Engineering, Infotech Oulu. Available at: http://herkules.oulu.fi/isbn9789526204048/isbn9789526204048.pdf. Accessed 1 Feb 2019
76. Wallbott HG (1998) Bodily Expression of Emotion. Eur J Soc Psychol 28(6):879–896
77. Wang Z, Ruan Q (2010) Facial expression based orthogonal local fisher discriminant analysis. In Proc. ICSP, 1358-1361
78. Westera W, Nadolski R, Hummel HGK, Wopereis I (2008) Serious Games for Higher Education: A Framework for Reducing Design Complexity. J Comput Assist Learn 24(5):420–432
79. Westera W, van der Vegt W, Bahreini K, Dascalu M, Van Lankveld G (2016) Software s for Serious Game Development. In: Connolly T, Boyle L (eds) Proceedings of the 10th European Conference on Games Based Learning 6-7 October 2016. ACPI, Paisley, pp 765–772
80. White KW (1999) The online teaching guide: a handbook of attitudes, strategies, and techniques for the virtual classroom (First Edition). Allyn & Bacon, Inc. Needham Heights. ISBN = 0205295312
81. Wright A (2009) Mining the web for feelings, not facts. An Article in New York Times. 23 of August 2009. Retrieved on 16 of December 2017. Available online at: www.nytimes.com/2009/08/24/technology/internet/24emotion.html?pagewanted=all&_r=0
82. Yang S, Bhanu B (2011) Facial expression recognition using emotion avatar image. IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011), 866-871. Santa Barbara. https://doi.org/10.1109/FG.2011.5771364
83. Zhang Z (1999) Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiment with a Multi-Layer Perceptron. Int J Pattern Recognit Artif Intell 13(6):893–911
84. Zhang Z (2012) Microsoft Kinect Sensor and its Effect. IEEE Multimedia 19(2):4–10. IEEE Computer Society Press Los Alamitos. https://doi.org/10.1109/MMUL.2012.24
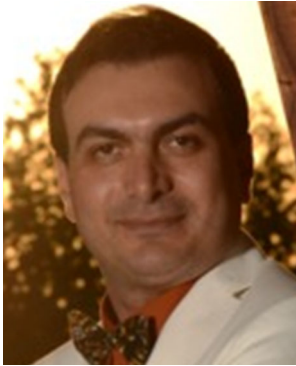
Kiavash Bahreini is a computer scientist with an interest in affective computing, data science, machine learning, real-time applications, and communication skills; he is an assistant professor at the Welten Institute, Research Centre for Learning, Teaching and Technology at the Open University of the Netherlands.

Wim van der Vegt is a software engineer with an interest in developing educational software, problem-solving, debugging software, computer interfacing, and performance improvements; he is a senior software engineer at the Welten Institute, Research Centre for Learning, Teaching and Technology at the Open University of the Netherlands.

Wim Westera is an educational media researcher with an interest in serious gaming and simulation; he is a full professor at the Welten Institute, Research Centre for Learning, Teaching and Technology at the Open University of the Netherlands.