Pilot study of real-time Emotional Recognition technology for Secondary school students

Cèlia Llurba, Gabriela Fretes, Ramon Palau

Universitat Rovira i Virgili, Departament Pedagogia. Crta Valls s/n,
43007 Tarragona, Spain
celia.llurba@urv.cat, gabriela.fretes@estudiants.urv.cat, ramon.palau@urv.cat

Abstract. The large variety of students in a class makes the teaching task complex, making it difficult for the teacher to personalise learning to each student. Since students should be at the centre of the educational process, it is necessary to know them better, so this study aims to explore the possibilities of using a camera for emotion recognition (ER) with a view to the potential use of this information to improve the teaching-learning process. To accomplish the aim it is previously necessary to develop and apply code capable of detecting faces, ER and transfer this data into a database for further analysis, which consists of establishing the first approximations to the relationship between students' emotions and other conditions (subject, time of day, academic performance). By monitoring the emotional state of students, if used properly, can improve educational processes, such as the teacher's decision-making in the classroom, as well as optimise attention to students, adjusting their methodology or focusing on a specific student.

Keywords: emotion recognition; learning process; teaching; database; artificial intelligence; learning analytics

1 Introduction

Teaching is a very complex, intellectually and emotionally demanding activity. Teachers have to deal with many students with increasingly heterogeneous profiles, as well as mastering their subject and teaching methodologies. Today, when it is so important that learners should be at the centre of the educational process, it is necessary to know them better. [1] Raised the importance of considering the people being taught, being sensitive to the characteristics of the students at the right time. Effective teachers have methods of getting to know their students' thoughts and feelings, which allows them to address their needs and can enhance the learning experience [2]. We believe that in this arduous task, technology can be an ally.

The challenges for teachers require the development of new strategies and tools to address them. In this pilot study, we proposed it would be useful to have artificial intelligence (AI)-based cameras to identify students' emotions in the classroom, process the data, and provide feedback to the teacher. Before widespread implementation in the educational environment, we must investigate the possibilities of this new technology from an educational point of view, to explore whether it is a useful resource for teachers and students, with the ultimate aim of improving the teaching and learning process. Specifically, having information on the emotional state of students in a systematic and objective way would allow teachers to make better decisions about their practice, such as planning and adjusting their methodology or catering for individual students. On the other hand, we must weigh up the difficulties or drawbacks, such as economic and ethical ones.

To accomplish this, we developed a code capable of detecting faces, captured by the digital camera Intel RealSense Depth Camera LiDAR L515), identifying and analysing their facial expressions to categorise them into primary emotions and transferring this data into a database for a subsequent statistical analysis. The analysis of the emotions data comprises establishing the first approximations to the relationship between students' emotions and other conditions (subject, time of day, academic performance), in order to explore the potential use of this information to improve the teaching-learning process.

A system that monitors students' emotions by detecting their facial expressions and how they interact with the classroom context is currently happening in Hangzhou No. 11 High School, in China. However, placing a camera with this emerging technology as a tool for educational improvement in a secondary school classroom is something new in our cultural context.

1.1 Background

To improve understanding of the educational potential of using emotional recognition cameras, it is useful to introduce several concepts and background information related to emerging technologies, classification of emotions and emotion recognition.

Emerging Technologies. In the current context, affected by the pandemic of COVID-19 pandemic causing restrictions on social interaction and even confinements, we are even more aware of the importance of technology in our daily lives [3]. Soon, classrooms could have integrated technologies beyond the use of devices. We could equip them with a photo or video camera that allows teachers to detect the emotions expressed in real-time by students and thus know how they relate to the teaching-learning process. This massive use of technology shows the impressive advances in recent decades and has become popular, when before it only existed in the most advanced industry. Therefore, artificial intelligence (AI), augmented reality (AR), Internet of Things (IoT), Big data, among others, are increasingly merging with everyday life, the so-called Fourth Industrial Revolution [4]. Thus, a large amount of data (big data) about people is collected from devices, which feed AI. In the same way we did in this, monitoring and collecting data from students through the camera device.

Emotions. [5], [6] argues that emotions are dynamic bodily dispositions that are at the basis of actions and that they found all human action on an emotion, thus emotions open up a space of possible actions to be performed. [7] The different emotions and thus different facial expressions are linked to distinct patterns of autonomic nervous system activity, which evoke a rather specific response in observers. Emotions can be divided as primary emotions, according to [8] into six different categories: happiness, sadness, surprise, disgust, anger and fear, plus neutral emotion. In fact, if we also consider secondary emotions, the ones that are reactive emotions to the primary emotions, then there are more complex emotions. The facial expression is the most important affective cue and correlates well with the body and voice [9]. The forehead and eyebrows, the eyes and eyelids and mouth are responsible for its manifestation on the face [10]. In this article we have used this classification as a basis for the creation of the code.

Emotions in academic context. Understanding students' emotions, especially during the classroom time, can help to improve the positive classroom emotional climate towards promoting academic achievement [11]. A recent systematic review [12] about academic emotions, students' emotions experienced in achievement and academic contexts [13], and the effects on learning, showed that, compared with negative academic emotions, positive academic emotions may be more effective at improving certain aspects of learning ef-fects, especially in high school and college students.

Emotion Recognition (ER). In the last decade, many studies have been conducted to understand the influence of emotions in education. There are several methods that help to make the global knowledge base measurable, as well as to ER, facial expression, posture, speech, from images or video frames [14], body language and skeleton motion, mobile device-based ER [15].

Having data such as the emotions of the students can help the teacher to make more accurate decisions about their teaching and attention to the student. There are models for decision making that emphasise the importance of having data. Data-driven decision making (DDDM) comprises systematic collection and use of many forms of data from a range of sources to improve student performance [16]; [17].

ER is important for the interaction between human and artificial intelligence, although there is a lot of data and label information for training [18]. It usually required a large amount of training dataset to obtain effective performance, with the design of the infrastructure of the system being the key to the effectiveness of student assessment [18].

2 Method

A pilot study was chosen because it allows us to approach a problem that has been less or not investigated at all. In our case, exploring whether emotional detection cameras for students has potential for the education process is a topic that has not been researched in our context.

The general aim is to explore the possibilities of using a camera for ER, with a view to the potential use of this information to improve the teaching-learning process. The first aim is to explore the emergence of students' emotions in the classroom context through the use of the camera facial recognition and emotional processing. The second aim is to explore the data provided and processed, for example, the relationship between students' emotions and subjects, students' emotions and time of day and students' emotions and their academic performance. Is the data obtained useful for the improvement of the teaching-learning process?

The methodology followed the User Centered Design (UCD), planning and analysing data, designing the sequence, developing, testing and revising. The participants were not randomly selected and were carried out in a natural classroom context. The study is a five-week longitudinal study, which is the most convenient way to capture students' emotions at various points in time. The first week was a test, so data was not counted for the statistical analysis. The camera was placed in two subjects, in Mathematics and Chemical classes.

To do this, it was necessary to carry out a previous task, the development of a code capable of detecting and identifying faces, analysing facial expressions and interpreting them as emotions.

The context of the classroom where the samples were taken and the data extracted with the camera, is a classroom with windows on one side, although there are also artificial light, fluorescent type; the blackboard is digital with white background. The classroom is of 23 students sitting at individual tables. At the time the samples were taken, the ambient temperature was between 18 and 24 degrees Celsius, it was the autumn of 2021. The classroom is not in the main building of the school, is located in the courtyard, in prefabricated barracks. During the lessons there is hardly any noise as in the classroom and at the adjacent classrooms the learners are high school students.

2.1 Resources

2.1.1 Programming Code

Below we will briefly describe the process of creating the code. The programming software chosen was Python, as computer scientists recommend it for new programmers because it is a powerful programming language, intuitive, with a clear syntax and works on Windows, Linux and IOS (Mac).

To facilitate programming, a code editor (IDEs) is needed to create and execute the code. Usually, an IDE is composed of a code editor, automation tools and a debugger. The IDE chosen was PyCharm, from the company JetBrains24. We used Azure Face for ER, but the face detection and identification was done with OpenCV, because the identifications with Azure gave us some errors.

We must access the Emotion API using the Microsoft Azure Emotion Recognition Tool. To connect to the Emotion API we needed a key and an endpoint, which are two identifiers provided by Azure in this process. Then uploaded a face photo and Azure analysed the emotions. The result was displayed in the most convenient format, in our case in an Excel file.

Using PyCharm, we could program a code in Python, incorporating the additional packages of the libraries to achieve the following tasks:

- To take pictures with the camera of several students.
- To take photos captured by the camera and put them in folders in the file directory, using Python.
- To identify the faces in each photograph, again using Python and the packages.
- Once the faces are known, each one must be assigned to a student.
- Then these pictures of each student's face are analysed on Azure.
- The Azure platform reports face emotion.
- Using Python, it saves the information to the directory.
- Finally, process and analyse the data obtained.

In order to be able to detect faces with the code we used the CascadeClassifier function of cv2, which implements a method called Haar Cascades. To identify faces we used the function through the Local Binary Patterns Histograms method.

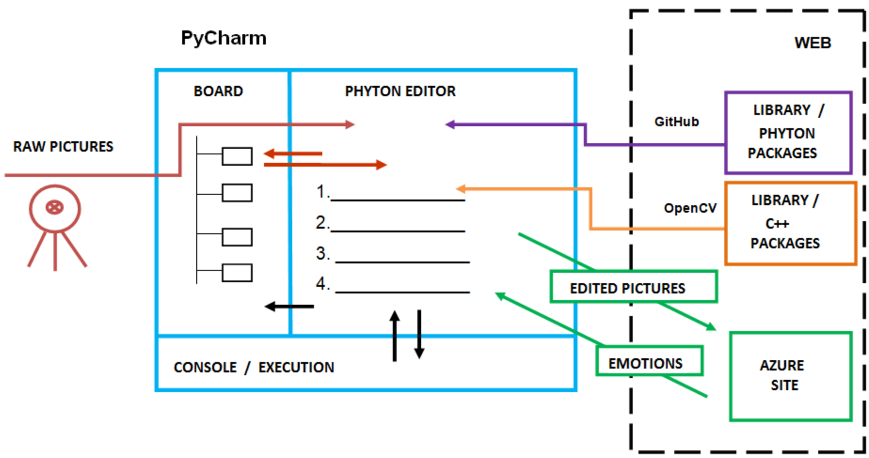

Fig. 1. Graphic. Workflow diagram

2.1.2 Devices

It needs a high performance camera with video recording capability connected to a laptop computer ready to run the required programs. The camera selected is an Intel RealSenseTM LiDAR L515. Its technical characteristics are:

- Depth capture from 0,25 to 9 m, up to 30 FPS5 at 1024*768 (XGA).6
- 2MP RGB7 camera

- Inertial Measurement Unit (IMU).8
- Image capture up to 30 FPS colour at 1920*1080 (FHD9)
- Class 1 laser

It is an Intel10 RealSenseTM LiDAR11 L515, high resolution, small and energy efficient. Its dimensions are 6,1 cm in diameter and 2,6 cm in height, with a weight of 100 g. It includes a tripod that makes it easy to place anywhere.

Fig. 2. Camera INTEL RealSense

RealSense is a set of 3D depth and acquisition cameras. These cameras can produce digital images and depth maps, using LiDAR technology. LiDAR is an optical technology that allows knowing the distance from a point to a surface or object by means of a point to a surface or object, by means of a fast scanning laser. It works by bouncing points of light off objects and measures the time it takes for the laser to travel back and forth from the camera to the object.

Intel claims that the Realsense L515 camera is the smallest laser telemetry device in the world. It has a very low power consumption, but is capable of a scanning range between 0,25 and 9 metres.

It creates a 3D map, or "dot plot" of everything that is around its sensor. This feature is not applicable for this study, since the objective is to take pictures and analyse them, not to make 3D images of the environment.

It has an internal processor that manages to capture moving objects with a minimum blur, which allows recognition of gestures and hand tracking. It has infrared ray vision, normal vision and night vision.

2.2 Experimental procedure

The camera was placed in the front row of the classroom, focusing directly on the two students. The first week was for adaptation to the camera and the photographs taken were discarded for analysis.

The camera detects students' faces and takes a picture when the face is detected, then identifies the face and matches it to a student. It must save these faces in personalised folders, for detecting the facial expressions of each photographed face, then relate the facial expression to the emotion.

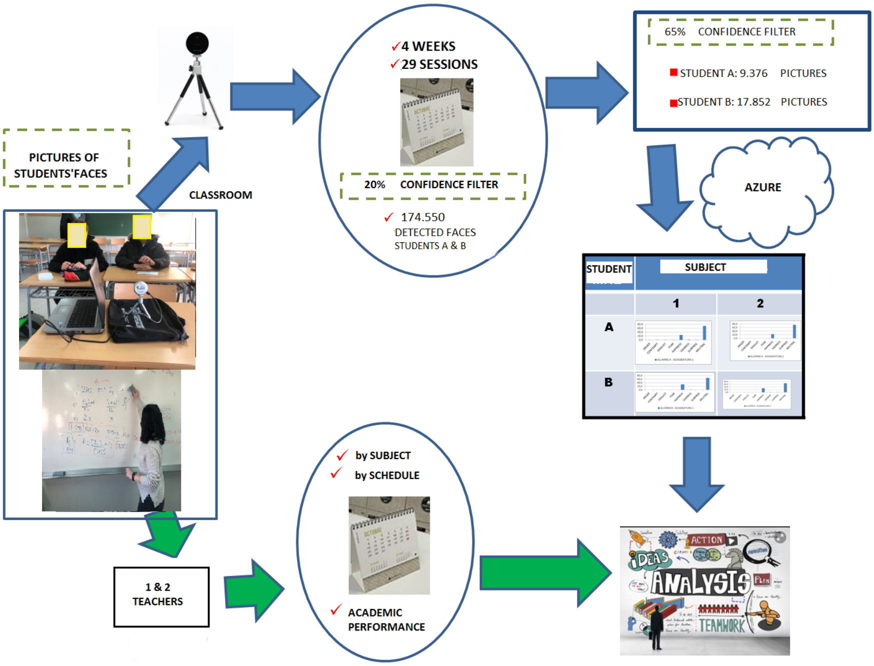

Fig. 3. Graphic. Process overview

2.3 Participants

Twenty-two high school students (non-compulsory secondary education) from a classroom at a Secondary school in Cambrils (Catalonia, Spain) volunteered to participate in the study. Because they are minors and the camera could capture their image within the classroom, the informed consent in writing was given to the family, the signature being essential to start the study.

Fig. 4. Students involved

The external conditions during the beginning of the experiment, mainly measured by the COVID-19, made that finally, the pilot study was carried out with only two students. They are both 17-year-old male students.

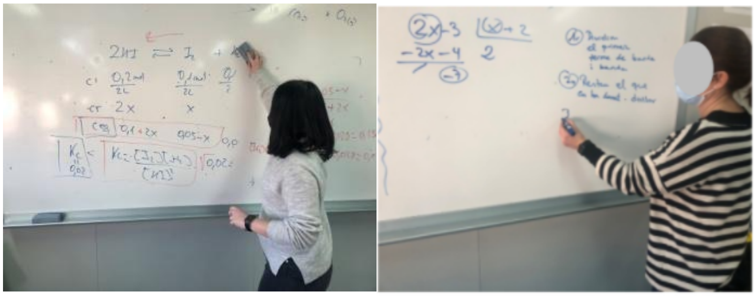

Fig. 5. Teachers involved

The teachers involved as volunteers were two females from the subjects of Mathematics and Chemical respectively. They were adequately informed and signed the informed consent.

2.4 Data analysis method

From the primary data obtained from the camera, we proceeded to perform statistical analysis using the Excel program and the open program JASP 0.16. We explored the data by making descriptive analysis, correlations among the different emotions, analysis of the emotions presented in each student in two different subjects (Maths and Chemical) and the time of day.

3 Results

The code was developed and worked correctly. It allowed us to detect and photograph faces with a digital camera in the classroom. Once the face is detected, it is identified with the corresponding pupil and stored in one of the two separate folders for each student. The emotion of each photographed face is then recognised. For example, results after the analysis of one picture:

Anger	Comptempt	Disgust	Fear	Happiness	Sadness	Surprise	Neutral
0	0	0	0	0	9,5	0	90,4

Fig. 6. Example of the result of the emotions' analysis of a picture.

Finally, the results can be viewed in Excel tables, which can be analysed, and then the second aim was achieved.

Next we will focus on the second aim, which is to explore the data obtained, 174.550 photographs have been taken over the four weeks (those of the test week are not included). These photographs are the ones taken by the camera in the classroom with a 20% confidence in the identification of the faces. Once the second filter was passed, set at 65% confidence, 27.209 were left, which are the ones that have been analysed to know the emotions.

Regarding the interpreted face images, captured throughout 28 sessions, student A has a total of 9.363 images (34,416%) and student B, 17.842 (65,584%). The difference in the number of identified photographs between the students is because student A will take part in three sessions less and for the different posture and movements of each one. There are sessions in which the quantity is similar (352 and 264) and others that are very different (67 and 939). In Mathematics there were 9.884 images (36,332%) and in Chemical 17.319 (63,668%).

The most commonly recognized emotion was neutral, followed by happiness in all of them. The neutral emotion is expected mostly in the classroom situation. What is to be evaluated is the recognition of characteristics of another emotion, although with a much lower percentage.

We chose to apply non-parametric tests because Levene's test was significant ($p<0,05$) suggesting a violation of the equal variance assumption, and the result of Kolmorov's test suggests non-normal distribution.

Relationships among emotions. Correlation bivariate analysis is performed between all emotion records. Spearman's correlations (see Table 1) showed significant relationships between almost all the different emotions, in the positive and negative sense, being of small or medium intensity. The strongest correlation was negative between happiness and neutral ($r=-8.848$, $p<0,001$), followed by the negative correlation between happiness and sadness ($r=-0,48$, $p<0,001$), the positive correlation between anger and disgust ($r=0.356$, $p<0,001$) and between fear and surprise ($r=0,35$, $p<0,001$). The remaining correlations were of small intensity.

Differences between students. The Mann-Whitney test indicated large U and significant differences ($p<0,001$) in all emotions, except fear, between the two students (see Table 2). The effect sizes were small.

Further analysis indicated that the median differences are extremely small according to the Hodges-Lehmann parameter (see Table 3). There were differences between medians in happiness, sadness and neutral. Thus, student A scored lower ($Mdn=0,100$) than student B ($Mdn=0,200$) on happiness ($U=7.922e+7$, $p<0,001$); student A scored higher ($Mdn=0,100$) than student B ($Mdn=0$) in sadness ($U=8.839e+7$, $p<0,001$); and in neutral emotion, student A scored lower ($Mdn=96,3$) than student B ($Mdn=95,7$), $U=8.590e+7$, $p<0,001$).

Table 1. Spearman's Correlations among emotions

Variable		Anger	Contempt	Disgust	Fear	Happiness	Sadness	Surprise	Neutral
1. Anger	Spearman's rho	-							
	p-value	-							
2. Contempt	Spearman's rho	0,206	-						
	p-value	< ,001	-						
3. Disgust	Spearman's rho	0,312	0,198	-					
	p-value	< ,001	< ,001	-					
4. Fear	Spearman's rho	0,18	0,076	0,316	-				
	p-value	< ,001	< ,001	< ,001	-				
5. Happiness	Spearman's rho	0,12	0,258	0,062	0,025	-			
	p-value	< ,001	< ,001	< ,001	0,016	-			
6. Sadness	Spearman's rho	0,317	0,081	0,133	0,111	0,504	-		
	p-value	< ,001	< ,001	< ,001	< ,001	< ,001	-		
7. Surprise	Spearman's rho	0,153	9,097E-04	0,138	0,312	0,121	0,067	-	
	p-value	< ,001	0,930	< ,001	< ,001	< ,001	< ,001	-	

Table 3. Descriptive Statistics about emotions

	Anger		Contempt		Disgust		Fear		Happiness		Sadness		Surprise		Neutral	
	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Valid	9363	17842	9363	17842	9363	17842	9363	17842	9363	17842	9363	17842	9363	17842	9363	17842
Missing	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Median	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,100	0,200	0,100	0,000	0,000	0,000	96,300	95,700
Minimum	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Maximum	67,900	69,400	83,100	67,600	23,700	4,300	1,600	2,400	100,000	100,000	94,700	96,400	98,300	99,900	100,000	100,000

Table 2. Independent samples

	W	p	Hodges-Lehmann Estimate	Rank-Biseral Correlation
Anger	9,021E+10	<,001	2,300E-05	0,08
Contempt	8,965E+07	<,001	4,680E-05	0,073
Disgust	8,405E+07	<,001	5,233E-05	0,006
Fear	8,347E+07	0,696	-2,040E-05	-7,252E-04
Happiness	7,922E+07	<,001	-4,327E-05	-0,052
Sadness	8,839E+07	<,001	5,475E-05	0,058
Surprise	8,705E+07	<,001	3,337E-05	0,042
Neutral	8,590E+07	<,001	5,634E-05	0,028

Note: For the Mann-Whitney test, effect size is given by the rank biseral correlation.

Differences between subjects. There are no significant differences in the variables of contempt, disgust, fear between the two subjects, Maths and Chemical. Statistically, there is a significant difference in the anger, happiness and surprise variables but when interpreted considering other parameters such as confidence intervals and effect size, we conclude that there is no useful significance. The sadness variable shows a significant difference (W=1277e+7, p<0,001) with a small effect size (0.184). (see Table 4). The sadness score in Mathematics was higher (M=1,71, SD=5,1) than in Chemical (M=0,663, SD=2,9).

Table 4. Emotions in Subjects (Maths & Chemical). Independent Samples T-Test

	W	p	Hodges-Lehmann Estimate	Rank-Biseral Correlation
Anger	1,145E+07	<,001	6,875E-05	0,061
Contempt	1,082E+07	0,764	6,529E-05	0,003
Disgust	1,079E+07	949	5,583E-05	1,809E-04
Fear	1,078E+07	0,662	-8,055E-05	-0,001
Happiness	1,050E+07	0,017	-3,760E-05	-0,027
Sadness	1,277E+07	<,001	6,562E-05	0,184
Surprise	9,894E+06	<,001	-2,252E-05	-0,083
Neutral	1,046E+07	0,010	-0,001	-0,031

Note: For the Mann-Whitney test, effect size is given by the rank biseral correlation.
Note: Mann-Whitney U test.

Emotions and time of the day

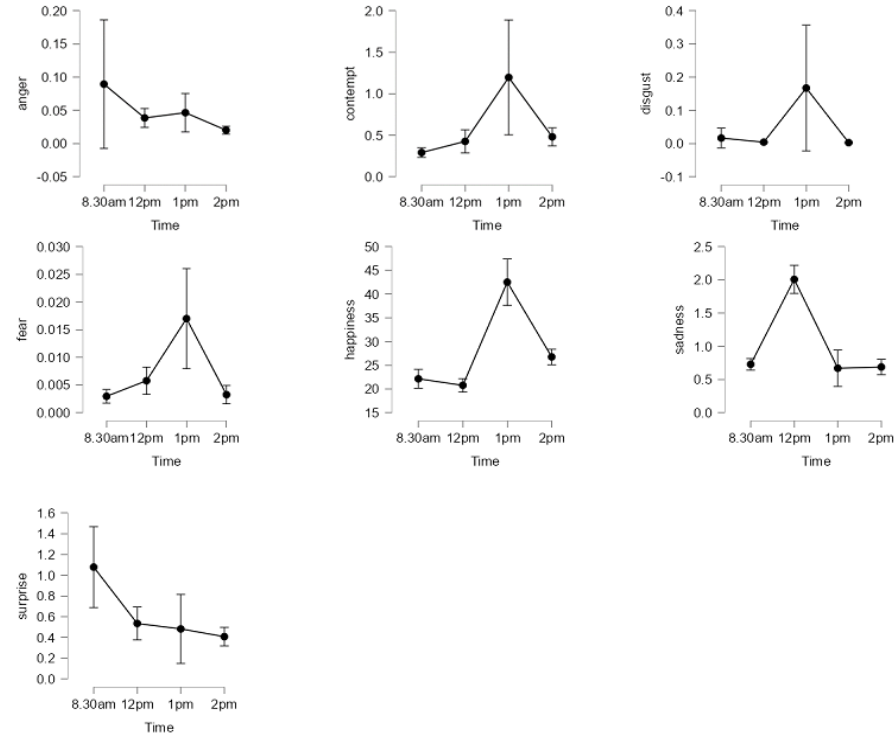

Fig. 7. Emotions during class time (Student A)

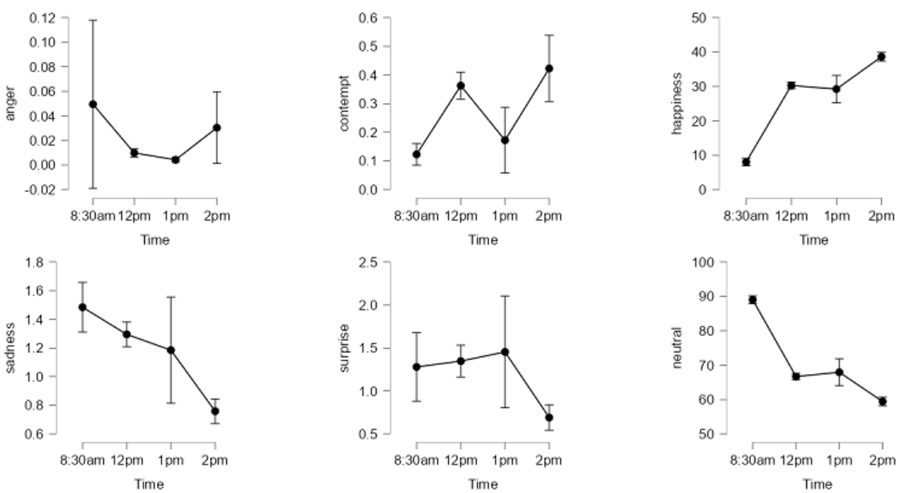

Fig. 8. Emotions during class time (Student B)

We obtained differences in the occurrence of all emotions according to class time (8:30am, 12pm, 1pm, 2pm). The graphical representation is presented in table 5 for student A, and table 6 for student B.

Anger. According to the Kruskal-Wallis test, there were significant differences in the anger variable by time of the day (H=26.536, p<0.001). The post hoc analysis, using the Bonferroni correction, indicated significant differences according to the time of day, specifically: between 8:30 am and 2 pm (<0.001), between 12 pm and 2 pm (p=0.003) and between 1 pm and 2 pm (p=0.004). The highest anger score occurred at 8:30am (M=0.089, SD=1.836), followed by 1pm (m=0.046, SD=0.276), 12pm (0.039, SD=0.382) and the lowest at 2pm (M=0.020, SD=0.147).

Student B also obtained significant differences in anger (H=27.333, p<0.001). Differences in anger were observed between 8:30am and 12pm (<0.001) and between 8:30am and 2pm (<0,001), being higher at 8:30am (M=0.049, SD=1.557) compared to 12pm (M=0.010, SD=0.151) and 2pm (M=0.030, SD=1.032).

Contempt. We found significant differences in contempt by time of day (H=13.519, p=0.004) according to the Kruskal-Wallis test. Dunn's post hoc analysis of comparisons, with Bonferroni correction, indicated differences between 8:30 am and 2 pm (p=0.003). The contempt score at 8:30am is lower (M=0.291, SD=1.087) than at 2 pm (M=0.481, SD=2.715). The contempt score at 8:30am is lower (M=0.291, SD=1.087) than at 2pm (M=0.481, SD=2.715).

In the same way, student B obtained significant differences in contempt (H=73.533, p<0.001). They were observed differences in almost all pairs of comparisons: between 8:30am and 12pm (p<0.001) and 2pm (p<0.001); between 12pm and 1pm and 2pm (both with p<0.001) and between 1pm and 2pm (<0.001). The highest score was at 2pm (M=0.422, SD=4091), followed by 12pm (M=0.363, S=2121), 1pm (M=0.172, SD=1266), the lowest is at 8:30am (M=0.123, SD=0.859).

Disgust. The Kruskal-Wallis test indicated significant differences in disgust by time of day (H=17.763, p<0.001). Dunn's post hoc Dunn's comparisons analysis indicates that there are differences by time of day between several pairs of times: between 8:30am and 1pm (p<0.001), between 12am and 2pm (p=0.001) and between 1pm and 2pm (p<0.001), with Bonferroni correction. The highest disgust score occurred at 1pm (M=0.167, SD=1.810), followed by 8:30am (M=0.017, SD=0.564), 12pm (M=0.004, SD=0.045) and 2pm (M=0.003, SD=0.029). In contrast, student B did not show significant differences for the disgust emotion across classes (H=3553, p=0.314).

Fear. According to the Kruskal-Wallis test, there were significant differences in the variable fear according to time of day (H=31.679, p<0.001). Dunn's comparison analysis, using the Bonferroni correction, indicated differences between various pairs of hours: between 8:30am and 1pm (p<0.001). and 1pm (p<0.001), between 12am and 1pm (p=0.001), and between 1pm and 2pm (p=0.001) and between 1pm and 2pm (p<0.001). The highest fear score was at 1 pm (M=0.017, SD=0.086), followed by

12pm (M=0.006, SD=0.066). At 2 pm and 8:30 am they have the same score (M=0.003). On the contrary, student B did not obtain significant differences for the fear emotion during the day (H=6.177, p=0.103).

Happiness. There were significant differences in the happiness variable according to the time of day (H=120.185, p<0.001) by the Kruskal-Wallis test. Dunn's post hoc analysis of comparisons with Bonferroni correction, indicated differences between all pairs of hours: between 8:30 am and 12 pm (p=0.004), 1 pm (p<0.001), and 2 pm (p=0.003); between 12 pm and 1 pm (p<0.001) and 2 pm (p<0.001); between 1 pm and 2 pm (p<0.001). The highest happiness score occurred at 1 pm (M=42.515, SD=46.833), followed by 2 pm (M=26.747, SD=41.083), 8:30 am (M=22.141, SD =38.054) and finally, the one at 12 pm (M=20.768, SD=37.684).

Student B also showed significant differences in happiness throughout class time (H=101.324, p<0.001). We observed differences between almost all the hours compared: between 8:30 and 12pm, 1pm and 2pm (all with p<0.001 level); between 12pm and 2pm (pz0.001) and between 1pm and 2pm (p<0.001). The highest score occurred at 2pm (M=38.627, SD=44.198) and the lowest at 8:30 a.m. (M=8,020, SD=25,173).

Sadness. According to the Kruskal-Wallis test, there were significant differences in the sadness variable by time of day (H=303.532, p<0.001). Dunn's post hoc analysis of comparisons, using the Bonferroni correction, indicated differences between all pairs of hours between 8:30 a.m. and 12 p.m. (p<0.001), 1 p.m. (p<0.001), and 2 p.m. pm (p<0.001); between 12 pm and 1 pm (p<0.001) and 2 pm (p<0.001); between 1 pm and 2 pm (p<0.001). The highest sadness score occurred at 12 pm (M=2.007, SD=5.676), followed by 8:30 am (M=0.728, SD=1.645), 1 pm (M=0.669, M =2.610) and, finally, the one at 2 pm (M=0.688, SD=2.879).

Student B also showed significant differences in the sadness emotion during the day (M=681.358, p<0.001). The highest score occurred at 8:30am (1,484, SD=3,948) and the lowest at 2pm (0.759, SD=3.053).

Neutral emotion. Kruskal-Wallis test indicates significant differences in the neutral emotion variable according to the time of day (H=96.126, p<0.001). Dunn's post hoc analysis, using Bonferroni, indicates that there are differences between 8:30 am and 12 pm (p<0.001), 1 pm (p=0.031); between 12 pm and 2 pm (p<0.001), and finally, between 1 pm and 2 pm (p=0.009). The highest neutral emotion score occurred at 8:30 am (M=1.078, SD=7.397), followed by 12 pm (M=0.535, SD=4.257), 1 pm (M=0.482 SD03.178) and 2 pm (M=0.408, SD=2.245).

Student B had significant differences in neutral emotion throughout the day (H=617.216, p<0.001). Differences were shown between 8:30am and 12pm, 1pm and 2pm (all p<0.001), between 12pm and 2pm (p<0.001) and between 1pm and 2pm (p<0.001). The hour with the most neutral score was 8:30am (M=59.444, SD=44.390).

Academic performance. Finally, regarding the relationship between academic performance and emotions in the two participants, with the data obtained, we cannot analyse this relationship and draw conclusions. We only point out that student A got a lower score for happiness, higher for sadness, and lower for neutral than student B. Student A got a lower score with a 6 (out of 10) in both Maths and Chemical subjects, while Student B got a 10 in Mathematics and a 9 in Chemical.

4 Discussion

An exploratory study was chosen because it allows us to approach a problem that has been less or not investigated at all. In our case, testing whether emotional detection cameras for students has potential for education is a topic that has not been researched in our context.

In this pilot study, we collected a considerable amount of data from photographs obtained by installing the camera in the classroom. Subsequently, the faces and facial expressions of the students were recognised and categorised into six emotions and neutral emotion, all with monitoring and knowing the emotional state helps the teacher to take this into account.

The percentage of processed photographs differed between the two students (student A=34.416% and student B= 65.584%) not only because student A had three fewer sessions but also because of movements and posture, such as partially covering the face. This shows the complexity of recording and interpreting emotions, the sensitivity of the camera and the need to invest a lot of time in the data collection process to make it valid and consistent. [19] shows that the work of deep learning-based algorithms highly depends on the size of the dataset; in general, we need a large amount of training dataset to perform effectively.

In relation to the analysis of the obtained data, we have been able to explore several aspects, such as the relationship between the presented emotions, comparison of the two students and the relationship between emotions and time of day. Regarding the relationship between the emotions (27205 records) coming from the two students, draw the expected relationships (positive relationship between emotions of the same sign, or negative relationship between opposite emotions). Regarding the differences between the students in Maths and Chemical, we observed that there are hardly any differences if we consider each emotion globally, with sadness being the only significant difference between the two students. We interpret this because emotions arise at a particular time and in a particular context, precisely something shared by both. In order to get a more accurate understanding, we should monitor subjects in different classrooms/times.

Regarding time and its effect on emotions, we found significant differences between times in each of the emotions. There are similarities and differences between the two participants depending on the time of day. We inform that the school timetable of the students is continuous, from 8:30 am to 3:00 pm. We found that at 8:30 am both felt more anger and more neutral emotion, and at 2 pm, less anger and less sadness. With this data, it is not possible to draw conclusions, although there is not a linear pattern but a curvilinear one, and what is possible is to get to know each individual student better. In addition, a study with the whole class group would allow a joint pattern to be established. It is important to carry out this type of study, because we have not found previous research that relates emotions and class time. There are studies that look at academic performance and school time, for example, empirical evidence of delaying the start time of the school day documented a positive impact on students' academic performance [20]. [21] Evaluated performance as a function of time of day using standardised tests and competencies, based on almost 2 million US students aged 11-16, and found that students learn more in the morning, especially in subjects such as Mathematics and Language.

In Spain, there are very few studies on the impact of educational schedules in general [22]. It is possible that this is because it is difficult to monitor and that emotions are changeable and more difficult to assess and record continuously. However, thanks to the camera and software used in this study, it will be easier to explore this issue in future research.

In relation to academic performance in our scan, the student with the most positive and neutral and least negative emotion scored excellent. This is in line with the literature, which agrees that positive emotions are related to better learning processes [23], while negative emotions have a negative effect [24], or may have ambivalent effects [12]. A future study would be necessary to establish correlations between emotions and academic performance, controlling for variables such as absenteeism.

5 Conclusion

We have been able to create a code capable of detecting and recognising faces, identifying emotions and converting images into data, which can be processed, analysed and stored. This has allowed us to establish the occurrence and frequency of seven emotions and the neutral emotion for each student. We know we cannot draw conclusions about the content obtained as it is a pilot study, but it is possible to glimpse the potential usefulness of having systematised emotional information made possible by ER technology together with contextual variables for the teacher, and even for other educational levels, and the benefit it would have for students. We see this pilot study as a first step in exploring the potential of ER in the classroom in our cultural context.

At present, teachers carry out their task, *per se* demanding, in a complex scenario. Secondary schools in Spain have a ratio of 25 students, compared to 23 and 21 in OECD and EU countries, respectively [25]. Besides the number of students, the heterogeneity of students makes it difficult to watch and offer personalised attention during the class session. Given this, applied AI can contribute to improving the teaching and learning process by providing feedback to the teacher. Knowing the students' emotions allows access to their genuine reactions to the learning environment. By detecting emotions that do not promote good learning, assuming that positive emotions improve the process, it is possible to intervene, by adjusting the learning activities or by focusing on a specific student.

In addition, it is possible to monitor the level of different emotions according to the time of day, as we have seen. This could help school principals and educational administration to evaluate and change timetables in a justified way to optimise students' well-being and academic performance.

With the stored data, long-term monitoring of the class group or individual students is possible, who may need special attention if a different pattern of emotional expression is detected. This opens up the possibility of preventing and detecting not only learning problems but also mental health problems.

Another contribution of our study is to propose a plausible alternative or complement to ER. Usually, emotions are self-reported through self-report and not in real-time, contributing to self-perception and memory biases.

Besides the face-to-face classroom, the use of ER can be useful in virtual classrooms, especially after the COVID-19 pandemic caused the explosion of virtualisation of synchronous classes with the use of cameras.

Finally, we provide evidence that it can bring a technological trend into the classroom with affordable financial resources and basic programming skills. This is in line with smart education, which embraces the technologies of the Fourth Industrial Revolution [26], which is the most revolutionary technological and social transformation, making things remote and increasing the efficiency and pleasure of our personal lives. This means that this kind of system could be extended to schools with no very large investment by the educational administration. A system like the one applied allows us to access data systematically and continuously, so we will get answers for which it would have been very difficult or impossible before.

Limitations. We highlight the inherent limitations of a pilot design, it is basically not possible to draw general conclusions and the sample is small.

About the limitations of the study and its development. On the one hand, health measures to prevent the spread of COVID-19, such as face masks, made it difficult to detect faces. We provided the students with transparent masks so that the camera could capture the mouth, essential for ER. However, this made it difficult to capture images when students were not facing the camera. On the other, the technical characteristics of the camera indicated a range of up to 9 metres. However, it had difficulty detecting beyond 3 metres. Thus, we conducted the exploratory study with fewer subjects than originally planned.

The observations were carried out on two male students, so it is not possible to explore gender differences, and they were carried out in the same learning context, so emotions may tend to be similar. Lastly, Mathematics and Chemical both require a high level of abstraction, which does not allow to compare the emotions generated with a different learning content.

Future research. On the basis of the results obtained, future lines of research can be considered.

- Identify more precisely the emotions expressed by students in response to specific learning content, methodologies, teaching style, and classroom environment, among other conditions.
- Monitor teachers' emotions, due to the importance of the teacher's role in the class climate, emotional development and student outcomes.
- Provide feedback to teachers and students on recorded emotions to assess their perception and usefulness for autoregulation.
- A study with a larger sample would be necessary to analyse the combined influence of emotions, academic performance and time of day.
- The evolution of the pilot study is to display the information to the teacher in real-time.

References

1. Jackson, P. W. (2002) Práctica de la enseñanza. Buenos Aires: Amorrortu.
2. Fisette, J. Getting to Know Your Students. Taylor & Francis Group , 81(7), pp. 42–49 (2013). https://doi.org/10.1080/07303084.2010.10598508
3. Palau, R., Fuentes, M., Mogas, J. & Cebrián, G.. Analysis of the implementation of teaching and learning processes at Catalan schools during the Covid-19 lockdown. Technology, Pedagogy and Education, 30(1) 183-199 (2021). https://doi.org/10.1080/1475939X.2020.1863855
4. Ross, P. & Maynard, K. Towards a 4th industrial revolution, Intelligent Buildings International, 13(3), 159-161. (2021). https://doi.org/10.1080/17508975.2021.1873625
5. Maturana, H. Emociones y lenguaje en educación y política. Dolmen. Santiago (1990).
6. Maturana, H. El sentido de lo humano. Hachette. Santiago (1992).
7. Ekman, P., Friesen W.V., & Ellsworth, P.. Emotion in the Human Face, ed., second ed. Cambridge Univ. Press (1982).
8. Ekman, P., O'Sullivan, M., Friesen, W.V & Scherer, K.R. Invited article: Face, voice, and body in detecting deceit. Journal of Nonverbal Behavior 15, 125–135 (1991). https://doi.org/10.1007/BF00998267
9. Zeng, Z., Pantic, M., Roisman, G.I & Huang, T.S. A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 1, pp. 39-58 (2009). https://doi.org/10.1109/TPAMI.2008.52
10. Caballo, V. E. Manual de evaluación y entrenamiento de las habilidades sociales. Madrid: Siglo XXI (1993).
11. Le-Quang, B. L., Dao, M. S., & Nazmudeen, M. S. H.. Wemotion: A System to Detect Emotion Using Wristbands and Smartphones. Communications in Computer and Information Science, 968, 92–103 (2018). https://doi.org/10.1007/978-981-13-5758-9_8
12. Tan, J., Mao, J., Jiang, Y., & Gao, M. The Influence of Academic Emotions on Learning Effects: A Systematic Review', International Journal of Environmental Research and Public Health 2021, Vol. 18, Page 9678. Multidisciplinary Digital Publishing Institute, 18(18), p. 9678 (2021). https://doi.org/10.3390/ijerph18189678
13. Pekrun, R. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. Educational Psychology Review, *18*(4), 315–341. (2006). https://doi.org/10.1007/s10648-006-9029-9
14. Hossain, M. S., & Muhammad, G. Emotion recognition using deep learning approach from audio–visual emotional big data. Information Fusion, 49, 69–78 (2019). https://doi.org/10.1016/j.inffus.2018.09.008

15. Boulton, H., Brown, D., Standen, P., Belmonte, M., Kwiatkowska, G., Hughes-Roberts, T., & Taheri, M. Multi-modalities in classroom learning environments. INTED2018 Proceedings, 1, 1542–1547 (2018). https://doi.org/10.21125/inted.2018.0264
16. Choppin, J. Data Use in Practice: Examples from the School Level Investigating Teaching Candidates' Experiences with Early edTPA Implementation View project Developing Principles for Mathematics Curriculum Design and Use in the Common Core Era (ERGO) View project. (2022).
17. Marsh, J. A., Pane, J. F., & Hamilton, L. S. Making Sense of Data-Driven Decision Making in Education: Evidence from Recent RAND Research. Occasional Paper. RAND Corporation (2006).
18. Choi, D. Y., & Song, B. C. Semi-Supervised Learning for Continuous Emotion Recognition Based on Metric Learning. IEEE Access, 8, 113443–113455 (2020). https://doi.org/10.1109/ACCESS.2020.3003125
19. Altrabsheh, N., Cocea, M., Fallahkhair, S. Learning Sentiment from Students' Feedback for Real-Time Interventions in Classrooms. International Conference on Adaptive and Intelligent Systems (2014). https://doi.org/10.1007/978-3-319-11298-5_5
20. Edwards, F. Early to rise? The effect of daily start times on academic performance.Economics of Education Review 31, 970-983 (2012). https://doi.org/10.1016/j.econedurev.2012.07.006
21. Pope, N.G. How the Time of Day Affects Productivity: Evidence from School Schedules. Review of Economics and Statistics 98:1, 1-11. (2016). https://doi.org/10.1162/REST_a_00525
22. Balart, P. Rendimiento y horario en las pruebas de conocimiento: evidencia de la Educación Superior. Fundación Ramón Areces y Fundación Europea Sociedad y Educación. (2021).
23. Martínez-Sierra, G., & García-González, M. Students' Emotions in the High School Mathematics Class: Appraisals in Terms of a Structure of Goals. International Journal of Science and Mathematics Education, 2(15), 349−369. (2017).
https://doi.org/10.1007/s10763-015-9698-2
24. Paoloni, P., & Vaja, A. Achievement emotions in assessment contexts: an exploratory study with university students. Innovación Educativa, 13(62), 135−159. (2013).
25. OECD, Education at a Glance 2021: OECD Indicators, OECD Publishing, Paris, (2021).
26. Mogas, J. & Palau, R. Smart Classrooms i l'adveniment de la Quarta Revolució Industrial: anàlisi dels factors clau per al disseny d'aules intel·ligents. [Doctoral dissertation, Universitat Rovira i Virgili] (2020).