International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I3P122

Volume 72 Issue 3, 249-259, March 2024

Original Article

Educational Quality Assessment System based on Emotions using Facial Images Applying Deep Learning Techniques

Victor Romero-Alva[1*], Sebastian Ramos-Cosi[1], Avid Roman-Gonzalez[1]

[1]*Image Processing Research Laboratory (INTI-Lab), University of Sciences and Humanities, Lima Peru.*

[*]*Corresponding Author: vromero@uch.edu.pe*

Received: 11 October 2024 Revised: 22 February 2024 Accepted: 29 February 2024 Published: 17 March 2024

Abstract *- In response to the shift towards online education prompted by the COVID-19 pandemic, this project aims to analyze students' emotional responses in a virtual classroom setting through the capture of facial expressions. By leveraging facial imagery, the initiative seeks to objectively assess levels of engagement, interest, or lack thereof, during online sessions in a discreet and non-intrusive manner. This methodology is intended to provide actionable insights to refine and enhance the effectiveness of future educational sessions, thereby optimizing the learning experience. Facial images are captured in real-time by students who have their cameras enabled during live class sessions. Subsequently, these images undergo processing to identify and evaluate the students' facial expressions. In its preliminary phase, the model successfully identifies and categorizes four primary emotions—happiness, sadness, anger, and surprise—with an impressive accuracy rate exceeding 90%, per the system's evaluation criteria. This accuracy is gauged based on the prevalence of each emotion throughout the class. The findings are then compiled and shared with educators as constructive feedback to inform and improve the planning and execution of subsequent sessions. Additionally, a Chrome browser extension has been developed to facilitate the deployment of this system within "Google Meet" platforms.*

Keywords *- Emotion detection, Recognition system, Educational quality, COVID-19, Virtual classes.*

1. Introduction

Due to the pandemic caused by the SARS-CoV-2 virus or simply COVID-19 (coronavirus disease) in the year 2020, which caused a serious global impact, almost all daily activities had to be suspended or modified without knowing exactly how, as the world was not prepared for a situation of such a virus infection, people suspected of infection had to be isolated in quarantine to avoid contagion and the spread of the virus [1]. This has also led to it being considered the most serious global health calamity of the century and has triggered a serious social and economic problem that continues to this day. By the year 2020, it was estimated that each month, there would be a decrease of up to 2% in annual GDP growth in the world's major economies, as most industries and economic sectors had to close, as well as most social sectors such as education [2]. On the other hand, using computer tools and ICTs has greatly intensified social development. It has led to information being used on a planetary scale, allowing a global digital transformation for society with an innovative approach due to the traffic of information and proliferation of the network, as they offer a great number of advantages and utilities [3]. This is why today, these digital tools can be considered almost indispensable for everyday life, all of this intensified even more in the face of an intransigent episode such as the global health crisis experienced by COVID-19 [4]. In this sense, one of the sectors most affected by the viral crisis has undoubtedly been the education sector, being practically the second one followed by the economic one, since, in mid-March 2020, more than 16 countries in Latin America had already suspended their face-to-face academic functions due to the difficult situation of universal emergency [5], so it was forced to immediately transfer all academic activity to a 100% online format, and this brought about that both teachers and students were involved in a quite drastic change and little used in regular teaching. Therefore, adapting to this new format has represented a real challenge for many teachers regarding reaching and understanding students due to several factors that were irrelevant before the COVID-19 context. In Brazil, in 2018, only 20% of teachers in that country participated in training in digital teaching methodologies and ICT [5]. Furthermore, it does not take a meticulous analysis to understand that teaching in an exclusively virtual format at a good level and with a high level of relevance is a challenge, especially for practical subjects. One of the direct and most evident consequences of this adaptation, now conceived as the new normal, and which is the product of inadequate training or prior preparation, is the stress experienced by teachers when they do not understand or know if they are reaching the student [6]. Similarly, due to the contingency, some types of professional burnout and mental overload may even develop, leading to manifestations of changing psychological behaviours [7]. This clearly represents a major problem in teaching, as, due to the lack of preventive strategies, the ability to achieve relevant and innovative competencies is only seen in a few teachers [8]. However, while it is true that adapting a virtual classroom as close as possible to a face-to-face classroom with the same level of impact and engagement is obviously complicated, it is also a reasonably necessary situation; therefore, ensuring maximum learning for the student population in each class is insistently essential.

For all of the above, it is considered that a solution contribution to improving the quality of virtual education lies in indirectly knowing the level of acceptance by the students present in each class, a feasible way to do this, taking into account that the common means of a virtual class for students is the provision of a video camera, is to use specialized algorithms to detect facial images during a video streaming to process them and determine the degree of interest / disinterest or emotional response by the viewer (student), in order to apply methods of correction and necessary modifications in teaching subsequently.

In this sense, the work presented in [9] exposes and highlights the technical capabilities of the fisherface algorithm compared to other techniques for face recognition characterization using facial image databases and GUI applications with the aim of developing an application. The technique of this algorithm uses a Principal Component Analysis (PCA) method followed by the use of the Fisher Linear Discriminant Linear (LLD) procedure. With these methods, the necessary image features are extracted. With the implementation of the face recognition system, a positive recognition response of 93% is obtained when the test image is different from the training image. On the other hand, a review of algorithms with applications more akin to the purpose of this work is presented in [10] and mentions that visual expressions are one of the main information paths in interpersonal communication, a key fact in the justification of the aim of the present work.

Both conventional and deep learning approaches to facial emotion recognition based on visual information are described. In this paper, we deduce increased facial emotion recognition performance using deep learning algorithms. Finally, the application of an emotion recognition system in real-time is developed in [11] having as a central tool a face detector through Key Points and a neural network applying Deep Learning, with the use of the OpenCV library of artificial vision since it is more efficient in response time, a key aspect in the execution of an algorithm that works in real-time. On the other hand, the development of the emotion detector model is carried out with the implementation of different neural networks with different layers and architectures and a support vector machine.

Consequently, this study is dedicated to identifying a pattern that reflects the quality of a class as perceived by students, which can be discerned through their facial features or emotional expressions. The primary goal is to design and deploy an emotion recognition system that analyzes facial images, utilizing deep learning techniques, within a videoconferencing context.This article unfolds as follows: in the second section, the methodology used in this work is presented, describing the different stages and processes. After this, the third section presents the results obtained from the implementation, and finally, the discussions and conclusions are presented.

2. Methodology

Having as a final proposal the complete development of a recognition and identification system concerning emotional characteristics, a previous study of the state of the art on the techniques and algorithms necessary for the present application is carried out. In this case, it is analyzed and determined that the import of the OpenCV and Mediapipe libraries is key to the realization of the projected algorithm, as they are able to work without problems in real time [12].

On the other hand, the two general stages of the system will be formed by the information capture component (video camera) and the information processing component (computer), these in turn are subdivided as shown in Figure 1. The stages mentioned above constitute the body of the emotion recognition system to be developed.

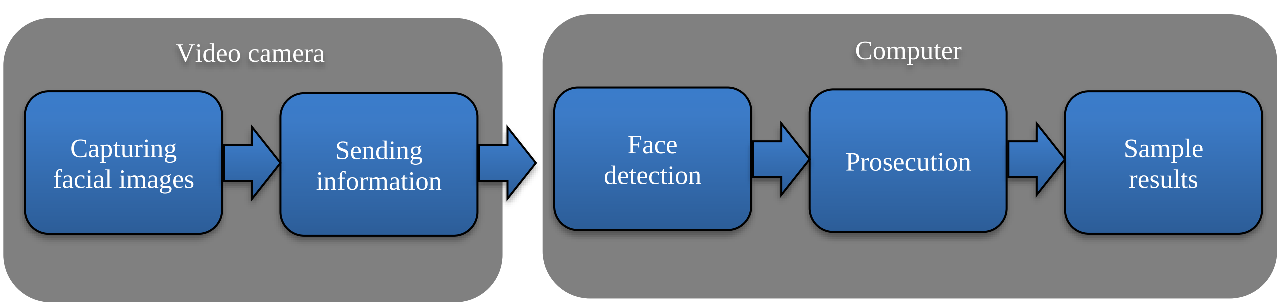

Fig. 1 Block diagram of the proposed system

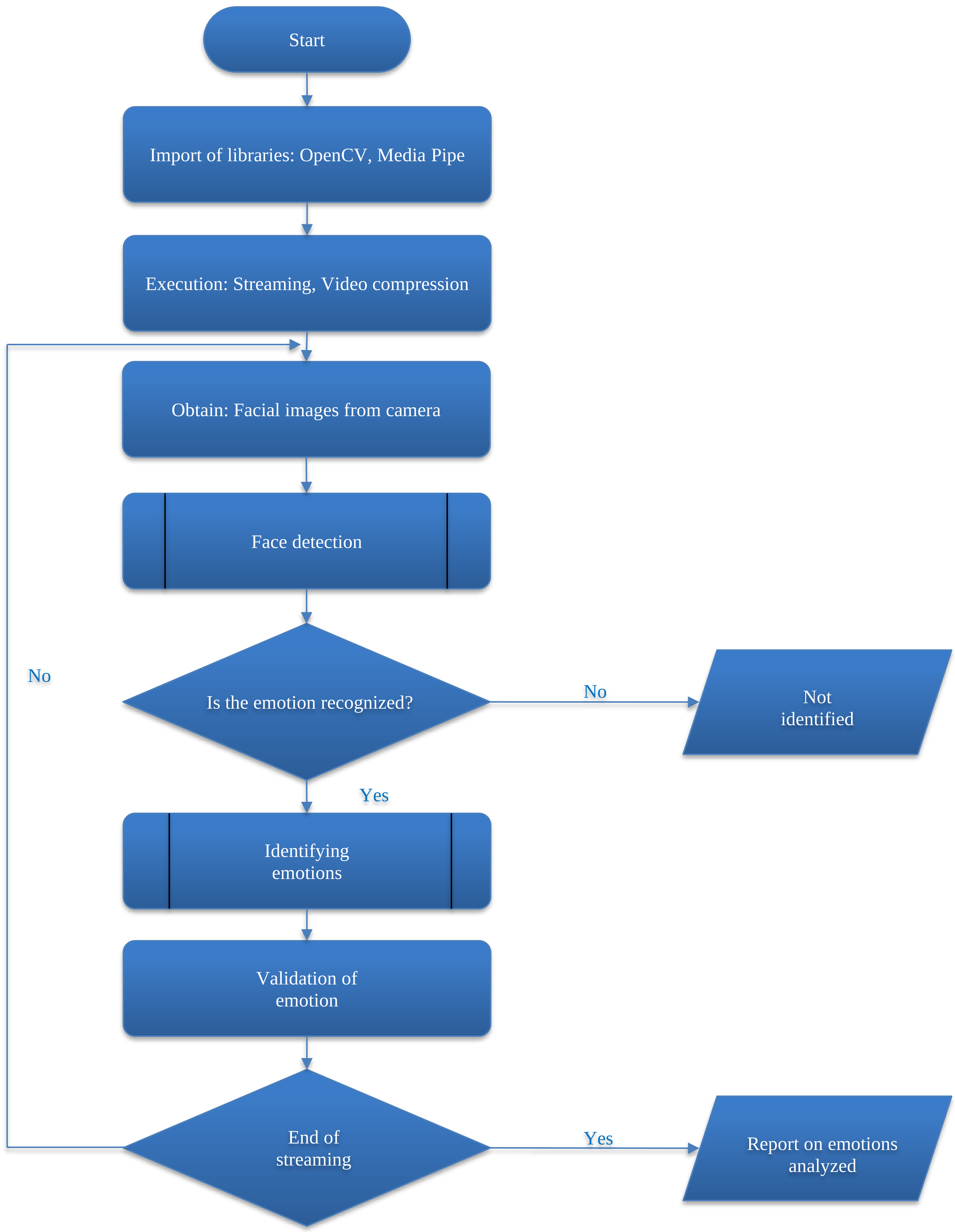

Fig. 2 Flowchart of the proposed system

2.1. Information Collection

This stage includes the first step for the entry procedure and access to the information carried out using a video camera as a hardware element, dividing into the acquisition and sending of a facial image or face. According to [13], this hardware element would also enter the preprocessing stage; however, for the present work, it is not considered within this stage as it is only used as an image input component and directly provides this information to the processing stage.

2.2. Information Processing

This section comprises the central stage of the emotional recognition system, essentially comprising the computer as the physical part and the algorithm as part of the software. This stage of the system receives the flow of information or data from the image-capturing device for the processing sequence itself [14], initially performing the required training of the algorithm developed in this stage and then analyzing and classifying according to emotionality criteria. In this sense, the algorithm concerning the identification of emotions contemplates the executed processes shown in the flowchart in Figure 2. Likewise, for the initial section of the process, the aforementioned requirements are considered, starting from the necessary libraries. The open-source OpenCV library provides access to the information captured through facial images and video [15] and, in turn, the link to the actual processing section.

A second necessary and facilitating resource that provides the fundamental tools for the training and detection of the evaluation and identification criteria in an image is MediaPipe [16]. This open-source software facilitates the detection of the displayed images by providing certain characteristics in the X, Y and Z coordinates and having the tracking function. It also allows presetting the image capture modes and the minimum confidence level of the image, thus providing the required input parameters to the program. Likewise, to work with deep learning techniques, it is necessary to define a language that is conducive and suitable for working with this type of learning; in this sense, Python is used as the primary language that works very well with MediaPipe. On the other hand, MediaPipe offers several resources to work on its software depending on the demanded or required application; among its portfolio of proposals mentioned, we have, on the one hand, the segmentation and detection solution and, on the other hand, the segmentation of selfies and the facial mesh [17], the latter two are focused on the analysis by face detection, which is of interest for the development of the proposed system.

2.3. Selfie Segmentation

Mediapipe's selfie segmentation tool [18] allows a person close to the camera to be separated or divided from the background around him or her by having an image or scene as input. In this way, a mask is obtained in which the regions where the lighter pixels are located (separated from the dark pixels) represent the people or subjects segmented in the capture.

2.4. Facial Mesh

This tool offers a large number of face detection alternatives, being highly flexible, to be applied in different environments (including mobile devices) and needs. It is based on the construction and arrangement of so-called "reference points" over the given region to be analyzed [19], which is an estimation of 468 points in 3 dimensions of the facial surface, whether it is a static image or a continuous streaming transmission such as video capture and transmission. It operates without the specific need for a depth sensor, thus using fewer processor resources and boosting its performance.

Both tools represent a good alternative to work on face detection and its extractable features. However, the most appropriate tools for the application of this particular work will have the attributes offered by the facial mesh implement since the aim is to analyze the facial images of the study subjects through a video transmission that may or may not be live. This tool employs very specific and detailed analysis resources focusing only on the face detected throughout a transmission of this type. The stages contemplated by the proposed face detection system handle the processing internally by extraction of evaluation criteria and seek to determine the emotion present in the facial image(s) of analysis. These stages of the system, which include both software and hardware, are strategically distributed, with the key moment being the central processing stage carried out by the core of the algorithm (Figure 2), which can also be visualized in Figure 3 in the preprocessing and information processing parts.

2.5. Validity and Reliability

On the other hand, the base parameter that provides the error percentage of any software is estimated with reliability [20]. In this sense, a reliability criterion is proposed for each emotion detected since a certain degree of detection or recognition accuracy is required for each emotion expressed. In this regard, Table 1 establishes margins for identifying a class of expressed emotion. Four basic emotions that are clearly differentiated and manifested will be analyzed; these correspond to the inherent reaction of the mood present in a person.

The ranges and scales mentioned in Table 1 correspond to the values required for detecting a certain emotion present. In this sense, considering that four emotions are considered for the analysis, they are divided by scale ranges to identify and classify them. In this regard, a total numerical range of 100 values is taken, and these are sectioned into groups of 25 to indicate their correspondence to a particular emotion. It is of relevant interest to know the level of accuracy required to take a certain action depending on the evaluation criteria, but consider a realistic margin of this level of accuracy [21]. Therefore, for the required accuracy and determination of the emotion, evaluation ranges concerning a pre-analysis of the detection are considered, which in turn provide preliminary information; these ranges are in the following percentages.

Table 1. Levels of emotion detection

Emotion	Scales and levels of accuracy		
	Ranks by scale	**Accuracy required**	**Difficulty of detection**
Cheerful/Happy	Range 0 – 25.	> 95%	Low
Sad/Discouraged	Range 25 – 50.	> 95%	Low
Angry	Range 50 – 75.	> 95%	Low
Amazed/Surprised	Range 75 – 100.	> 95%	Low

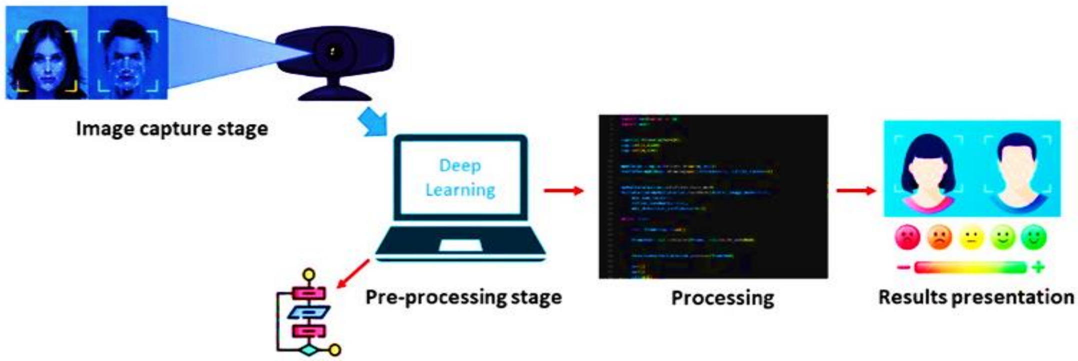

Fig. 3 System architecture

- Low accuracy: 10% – 45%
- Medium accuracy: 45% – 75%
- High precision: 80% – 99%

For the application of the proposed system, the aim is to obtain a high accuracy within the margins above 90%; ideally, as mentioned in Table 1, accuracies above 95% in detection are required. Being a system that uses deep learning tools for its processing, it is expected that it can learn itself in a short period of time and that the accuracy can be constantly improved [22][23].

2.6. Proposed System Architecture

In this sense, it can be understood that the general functioning of the system contemplates different stages and processes and that these correspond to elements present in the software and hardware [24]. In this respect, the proposed conceptual model or proposed architecture responds directly to the needs of the system itself. The representation of this model is shown in Figure 3. For the stage corresponding to the image or video capture, a Logitech Brio 90 4K webcam is considered for testing and experimentation, as it offers high resolution in image acquisition. The webcam has different resolution options being 4K at 30 frames per second, 1080p at 30 or 60 frames per second, and finally 720p at 30, 60 or 90 frames per second [25], and the processing equipment consists of an MSI laptop with a Ryzen 7 processor.

3. Results

The developed system first goes through a verification process by image types or identification stages with their corresponding previous configuration; in the first stage, the resources necessary for the application of the algorithm are checked; as part of the testing and verification process, the detection is checked on a static facial image entered in PNG format (Figure 4). The aim is for the algorithm to be able to arrange and align the key points on the identified facial image. As can be seen in the test image in Fig. 4, the reference points are correctly displayed and located over the entire surface of the detected test face, covering mainly eyes, nose and mouth in the same facial image. These correspond to the 468 points covered by the mesh. Compared to the static test image without real-time transmission, the detection and meshing behaves normally and does not represent any major difficulty in its deployment. Due to its versatility and lightness, the implementation of the programme code together with the necessary libraries is carried out and configured using the Visual Studio Code editor and some of the special tools for its rapid execution. The emotional correlation matrix in Figure 5 reveals how various detected emotions (Happy, Sad, Angry, Amazed) are associated with each other through correlation coefficients ranging from -1 to 1. A coefficient close to 1 indicates a strong positive correlation, implying that an increase in one emotion leads to a similar increase in another. In contrast, a coefficient close to -1 reflects a strong negative correlation, indicating that an increase in one emotion causes a decrease in another. A value around 0 indicates the absence of a significant linear relationship between the emotions evaluated. On the other hand, to start the execution of the system with video detection or streaming in real-time, the aforementioned Logitech Brio 4K camera is used as an image-capturing element, which provides the visual information of 13 megapixels with an FPS rate of 30. This allows the system to achieve detection in the streaming video with optimal tracking and positioning speed.

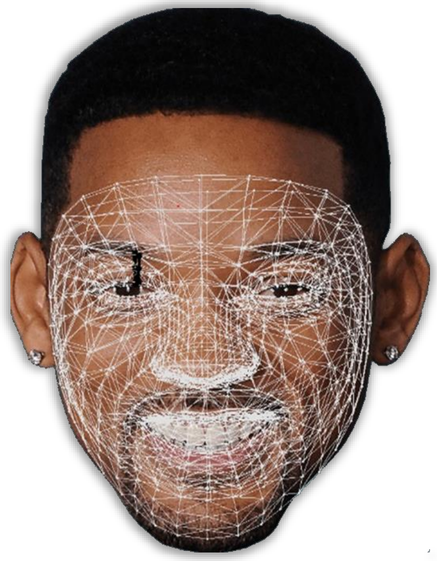

Fig. 4 Detected facial image with key points

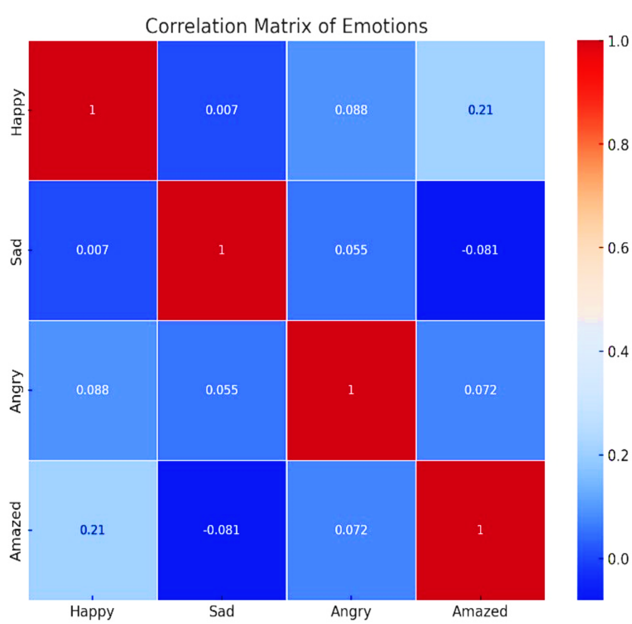

Fig. 5 The emotional correlation matrix

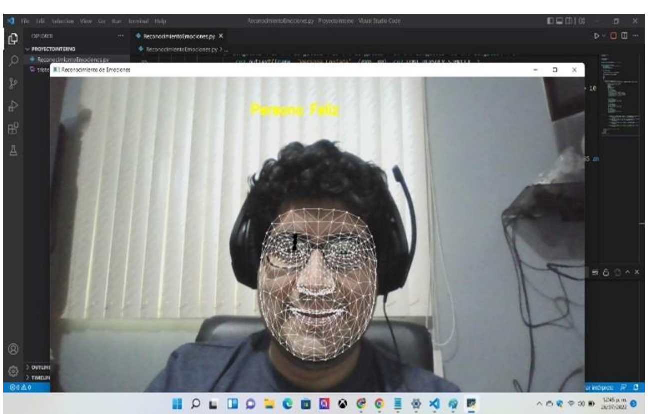

Fig. 6 Facial image detected with "happy" emotion

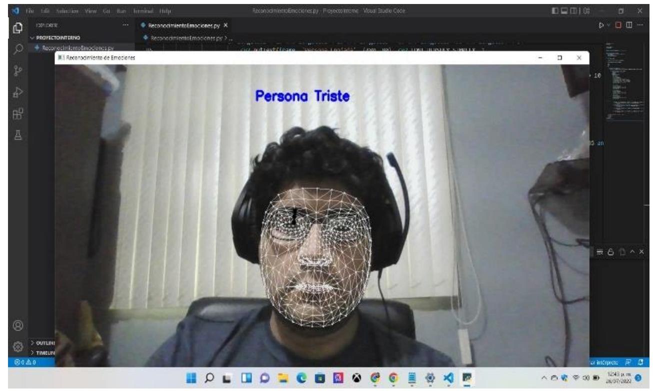

Fig. 7 Facial image detected with "sad" emotion.

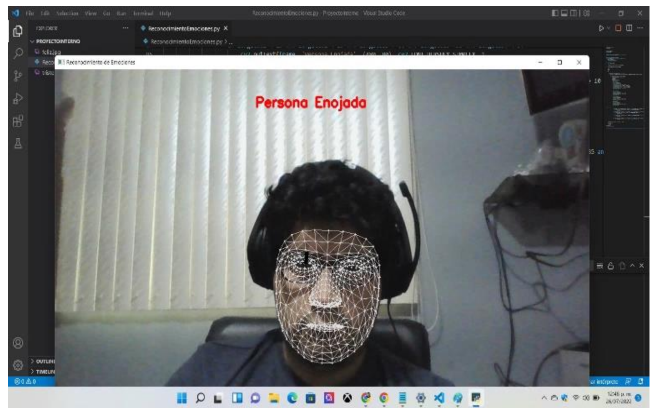

Fig. 8 Facial image detected with "angry" emotion

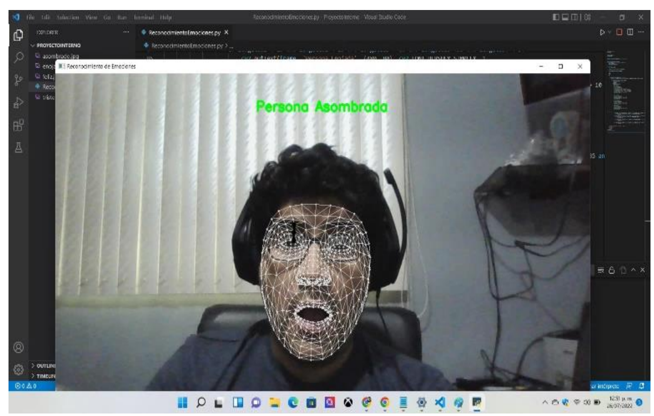

Fig. 9 Facial image detected with "amazed" emotion

The emotions detected and classified by the different reactions and facial movements made by the user are exposed as part of the tests, the emotions captured are four (happy, sad, angry and amazed).

Regarding the system's first interaction with the image capture, it can be seen in Figure 6 that it manages to detect the emotion corresponding to a "Happy" state by measuring the criteria established in the mesh implemented on the face and in the information points. The result is shown textually in yellow. Continuing in the same streaming with the corresponding tracking in progress, when a new change in the facial reaction occurs, a new emotion is detected, as shown in Figure 7; this emotion corresponds to “sad”; in this case, the detection result is shown at the top in blue. On the other hand, when the facial features change again and thus the arrangement of dots within the facial mesh, the emotion “angry” is detected, and the same detected emotion is displayed in red (Figure 8). Finally, the last emotion recognizable by the system is detected; this is the emotion of “amazed”, shown in Figure 9, which shows the result textually in green.

Having these four emotions detected and classified accordingly, it is analyzed that the correct detection is directly related to the measurement criteria between one point and another of the 468 included. In this sense, when the eyebrows are raised and the lip line is elongated, a “happy” emotion is detected while the eyebrows are furrowed. The lip line is short; that is, if the measurement between the points of the lips is smaller, a state of “angry” occurs; on the other hand, if the expression shows a short distance between the points of the horizontal line of the lips and with a slight fall and relaxation in the eyebrows, the system detects an emotion of “sadness”. Finally, when there is an expression that includes raising both eyebrows together with a separation between the middle part of the centre line of the lips, the system detects an emotion of surprise or “amazed”. In this way, the first verification of performance in real-time detection is carried out by means of a video transmission. All of this takes into account the detection levels of each emotion (Table I). However, as it is a software system, it is necessary to validate its operation during a prudential transmission time, and this validation is carried out by recording the moments when a correct detection is made and the moments when it is not made.

For this purpose, as an initial phase, constant 10-minute video transmission and capture are maintained as part of a period of error verification in a fluid transmission, randomly expressing the four possible emotions. For the evaluation phase of each recognizable emotion selected, and after the first 10 minutes of test transmission, a measurement criterion of 2 minutes was proposed for the final transmission. During this time, it was divided into 15 facial reaction moments of 8 seconds each, as shown in Figure. 10, from which it can be seen that, in general, there was an 80% detection rate of the emotion “happy”, with two clear occasions of erroneous detection in the same 15 moments, as they did not correspond to the real emotion, also, if applicable, a maximum value of 25.

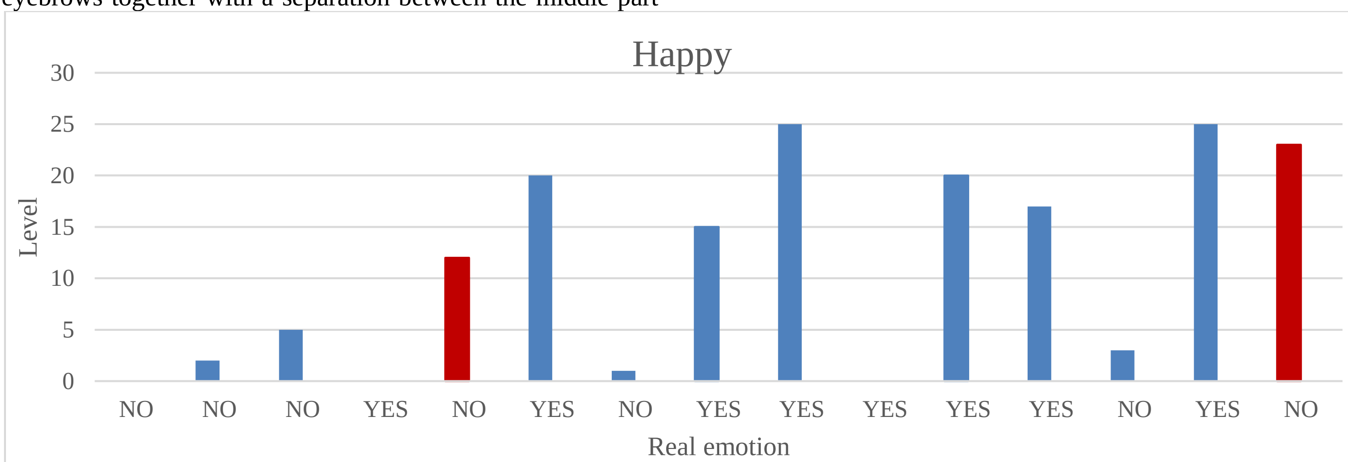

Fig. 10 Emotion “happy” detected in values from 0 to 25

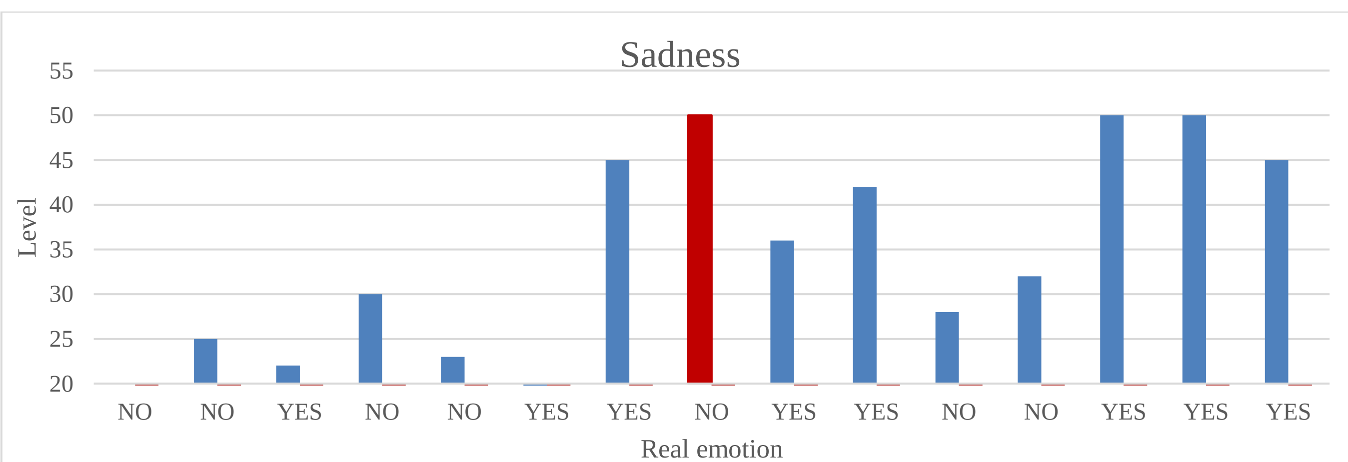

Fig. 11 Emotion “sadness” detected in values from 20 to 50.

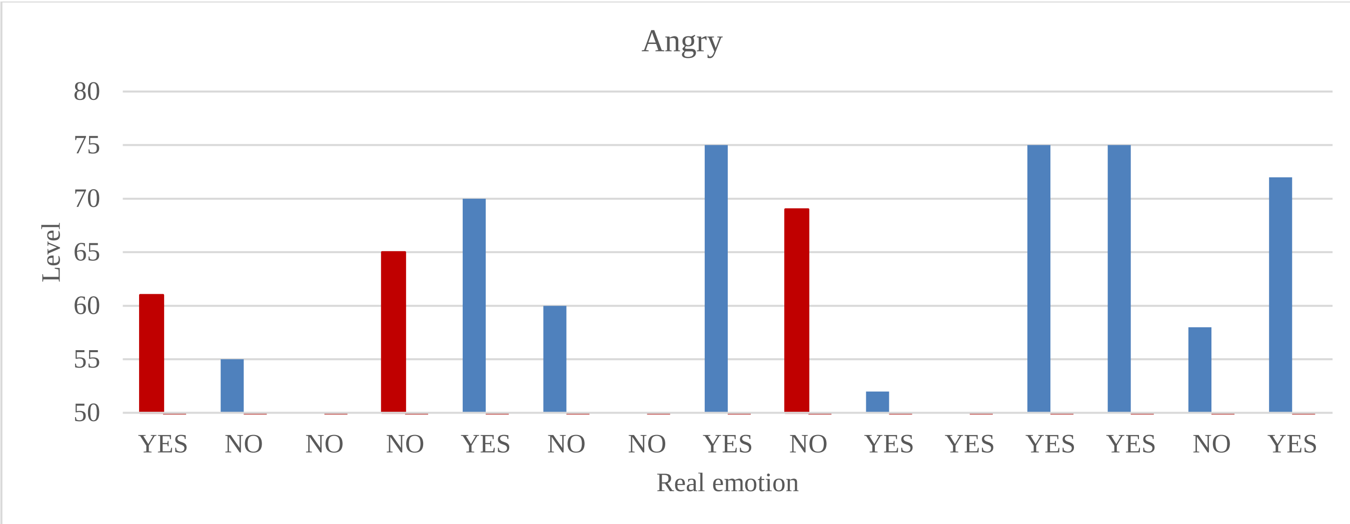

Fig. 12 Emotion "angry" was detected in values from 50 to 75.

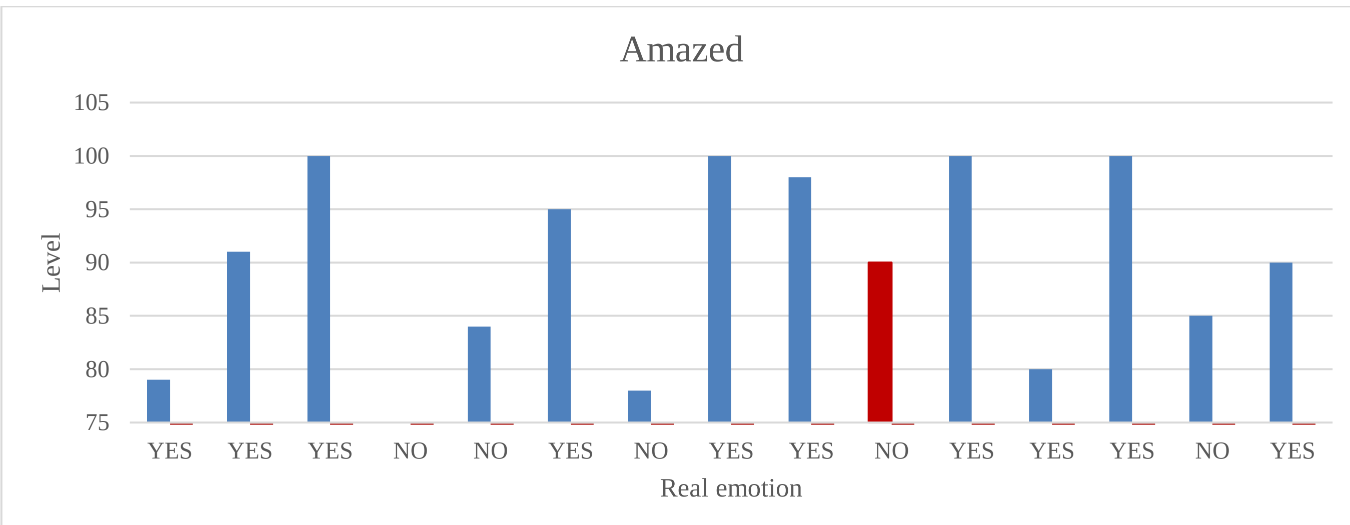

Fig. 13 Emotion "amazed" detected in values from 75 to 100.

Regarding the evaluation of the detection of the emotion "sadness", an 87% general detection was recorded, of which one occasion of erroneous detection was obtained as it was not a real emotion other than the one manifested in the transmission, as shown in Figure 11. For the precision of this emotion, values were taken in the range of 25 to 50.

In this evaluation, five clear moments were obtained in which level 40 of detection was exceeded, within which two recorded the highest value.On the other hand, evaluating the system's response to the emotion "angry" yielded an overall detection of 80%, with 20% not detected by the system. Of those obtained, 3 occasions of detection that exceeded the value of 60 were erroneous, as shown in Fig. 12. On the other hand, there were 5 occasions that had a detection value of 70 or more.Finally, in evaluating the response to the emotion "surprised", a 93% overall detection rate was obtained, with only one occasion in which there was no detection. Likewise, out of the 15 moments, 6 occasions of detection were counted with the highest level of accuracy (Fig. 13), and only one relevant erroneous occasion of value 90 out of the 4 detected. The range of values for this emotion remained between 75 and 100.

Having carried out the corresponding verification and evaluation of the four possible emotions to be detected, it is possible to proceed to implement a utility to link it to any transmission in progress coming from the different streaming platforms for the particular application of this work, the transmission of the Google "Meet" application is considered, for this, an extension is generated in the Chrome browser that allows direct access to these transmissions so that the execution of the recognition system in the current meetings in progress is more practical. This extension is represented by an icon with the University of Sciences and Humanities logo, as shown in Fig. 14 at the top right of the browser, indicated by a green box.

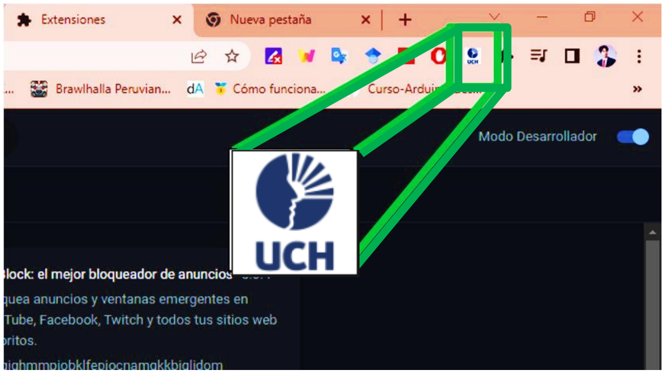

Fig. 14 Extension generated in the chrome browser

4. Discussions

The results obtained show the implementation of a unified system that manages to detect and identify four independent emotions present during a real-time video transmission; such a system can be portable and capable of being anchored as a browser extension (Chrome).

Corresponding "happy" to the first recognizable emotion according to the evaluation process carried out, Fig. 10 shows the 15 moments of response by the system to the emotion; 3 possible cases of response are analyzed, in which a successful detection occurs, in which an incorrect detection occurs and in which no detection occurs, it can be seen that the trend is towards a successful detection; however there are 3 occasions of an erroneous detection by having a value of more than 5, being 2 occasions the most critical since their values reached 23 out of 25. On the other hand, of the three moments where there was no detection, two corresponded to a real emotion that was not detected, which is assumed to be due to the distance between the camera and the user.

On the side of "sadness" as the second emotion evaluated, it can be seen in Fig. 11 that there were two maximum possible situations of successful detection out of 6, and only one maximum wrong detection; the other four false detections that did not represent a real emotion of sadness have less relevance according to the values achieved. It is analyzed that when the two minutes of transmission start to run, there is always a transition period in the first moments, so the values are not immediate maxima but have a small delay. Also, one occasion of real emotion was not detected, so no assessment occurred. One of the erroneous detections due to the lack of real emotion is related to the distance present in the points located in the area of the lips. For this reason, a correction was made to the measurement of the points arranged in the facial mesh.

Regarding the evaluations carried out on the emotions "angry" and surprised" (Fig. 12 and 13), there were slight differences that need to be analyzed; on the "angry" side, there were three clear occasions of erroneous detection due to a real false detection; however, the occasions of maximum correct detection reached 3 in this case, and another 2 with slightly less relevance.

On the other hand, there were occasions where no detection increased, these being 3, but only one represented a real emotion, so acceptable precision was demonstrated. Now, regarding the evaluation of the emotion "surprised", it is evident that for the system, it has been one of the easiest to recognize since the opening of the mouth and the raising of the eyebrows provide quite conclusive information; therefore, There was only one erroneous detection of medium relevance as it was not a real emotion. However, there were 7 relevant correct occasions that guaranteed the system's detection reliability. With the information provided, an emotional recognition report is displayed and sent in real-time based on the four detectable emotions. Likewise, a backup report could be generated and stored independently for its respective post-video streaming analysis.

5. Conclusion

This work started with the purpose of developing an emotion recognition system applying Deep Learning techniques that are capable of detecting four base emotions during video transmission. In this regard, the system has been able to respond satisfactorily with what was indicated, thus detecting a user's mood of "happy", "sad", "angry", and "surprised" the detection efficiency improved by up to 20%; after applying some corrections previously in view of the results and emotion detection evaluations obtained. Likewise, the extension generated in the Chrome browser has turned out to be partially functional due to some version incompatibilities and response speed. It is expected that this system can be scalable and, therefore, implemented on different video streaming platforms.

The system, in its entirety, has maintained an average emotion detection accuracy of over 90%, which validates its performance during its execution. The layout of the meshing presents a good level of response in tracking the detected face or facial image, together with the different measurement criteria provided by the points of interest, which are fundamental for correct detection.

Although it is true that the system still represents a pilot model and can be scalable and improved for more optimal training with the inclusion of more emotions, in general, for the first version of the model, the system behaves with acceptable and adaptable performance. On the other hand, with the system implemented and running in real-time using the different cameras available, the necessary information provided by the recognition system itself, which in turn comes from the students, is key to an adequate and objective self-evaluation of the teaching in each session and therefore helps to improve the virtual class sessions taught by the teachers.

References

[1] Ashok Kumar Verma, and Sadguru Prakash, "Impact of COVID-19 on Environment and Society," *Journal of Global Biosciences*, vol. 9, no. 5, pp. 7352-7363, 2020. [Google Scholar] [Publisher Link]

[2] Indranil Chakraborty, and Prasenjit Maity, "COVID-19 Outbreak: Migration, Effects on Society, Global Environment and Prevention," *Science of The Total Environment*, vol. 728, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Juan Fernando Grajales Escobar, and Yadir Milena Osorno Mira, "Globalization and the Importance of ICT in Social Development," *Reflections and Knowledge Magazine*, no. 11, pp. 2-9, 2019. [Google Scholar] [Publisher Link]

[4] Natalia García Fernández, María Luisa Rivero Moreno, and José Ricis Guerra, "Digital Divide in Times of COVID-19," *Revista Educativa Hekademos*, no. 28, pp. 76-85, 2020. [Google Scholar] [Publisher Link]

[5] Wildebeest. Eclac, *Latin America and the Caribbean in the Face of the COVID-19 Pandemic: Economic and Social Effects*, COVID-19 Special Report, pp. 1-15, 2020. [Google Scholar] [Publisher Link]

[6] Juan Luis Cortés Rojas, "Teacher Stress in Times of Pandemic," *Contemporary Dilemmas: Education, Politics and Values*, vol. 8, no. 1, pp. 1-11, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Gladys Galvis López et al., "Tensions and Realities of University Teachers in the Face of the Covid-19 Pandemic," *European Journal of Health Research*, vol. 7, no. 1, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Josnel Martínez-Garcés, and Jacqueline Garcés-Fuenmayor, "Digital Teaching Skills and the Challenge of Virtual Education Derived from Covid-19," *Education and humanism*, vol. 22, no. 39, pp. 1-16, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Mustamin Anggo, and La Arapu, "Face Recognition Using Fisherface Method," *Journal of Physics: Conference Series*, pp. 1-10, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Byoung Chul Ko, "A Brief Review of Facial Emotion Recognition Based on Visual Information," *Sensors*, vol. 18, no. 2, pp. 1-20, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[11] Héctor Ribes Gil, "*Development of a Facial Emotion Recognition System in Real Time*," Final Degree Thesis, Autonomous University of Barcelona, School of Engineering, pp. 1-9, 2017. [Google Scholar] [Publisher Link]

[12] Yejin Kwon, and Dongho Kim, "Real-Time Workout Posture Correction Using OpenCV and MediaPipe," *Journal of KIIT*, vol. 20, no. 1, pp. 199-208, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Diego Fabián Centurión Talavera, and Carlos Domingo Almeida Delgado, "Facial Recognition using Artificial Neural Networks on Raspberry Pi," *FPUNE Scientific*, no. 16, pp. 17-24, 2022, [Google Scholar] [Publisher Link]

[14] Juan José Gutiérrez Leguizamón et al., "Colombian Sign Language Recognition Using Convolutional Neural Networks and Motion Capture," *Technurae*, vol. 26, no. 74, pp. 70-86, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Ivan Culjak et al., "A Brief Introduction to OpenCV," *2012 Proceedings of the 35th International Convention MIPRO*, Opatija, Croatia, pp. 1725-1730, 2012. [Google Scholar] [Publisher Link]

[16] L. León et al., "Development of a Face Detection System with or without a Mask using Software," *20th LACCEI International Multi-Conference for Engineering, Education, and Technology, Education, Research and Leadership in Post-pandemic Engineering: Resilient, Inclusive and Sustainable Actions*, Boca Raton, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Giuseppe Massimo Bernava et al., "An Advanced Tool for Semi-automatic Annotation for Early Screening of Neurodevelopmental Disorders," *Image Analysis and Processing*, *ICIAP 2022 Workshops*, Lecce, Italy, pp. 154-164, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Arya Khanorkar et al., "Background Removal of Video in Realtime," *3C ICT: Applied ICT Development Notebooks*, vol. 11, no. 2, pp. 195-206, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Mukhriddin Mukhiddinov et al., "Masked Face Emotion Recognition Based on Facial Landmarks and Deep Learning Approaches for Visually Impaired People," *Sensors*, vol. 23, no. 3, pp. 1-23, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Andrés Jaramillo Ortiz, M. Robinson Jiménez, and Olga Lucia Ramos, "Quality Inspection for an Industrial Production System Based on Image Processing," *Investigation*, vol. 18, no. 41, pp. 76-90, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[21] Todd M. Dewey et al., "Reliability of Risk Algorithms in Predicting Early and Late Operative Outcomes in High-Risk Patients Undergoing Aortic Valve Replacement," *The Journal of Thoracic and Cardiovascular Surgery*, vol. 135, no. 1, pp. 180-187, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[22] Mackenzie Weygandt Mathis, and Alexander Mathis, "Deep Learning Tools for the Measurement of Animal Behavior in Neuroscience," *Current Opinion in Neurobiology*, vol. 60, pp. 1-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Alison J. Fairbrass et al., "CityNet-Deep Learning Tools for Urban Ecoacoustic Assessment," *Methods in Ecology and Evolution*, vol. 10, no. 2, pp. 186-197, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[24] John Klein, and Hans Van Vliet, "A Systematic Review of System-of-Systems Architecture Research," *Proceedings of the 9th International ACM Sigsoft Conference on Quality of Software Architectures*, pp. 13-22, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[25] Terisara Micaraseth et al., "Coffee Bean Inspection Machine with Deep Learning Classification," *International Conference on Electrical, Computer, Communications and Mechatronics Engineering*, Maldives, pp. 1-5, 2022, [CrossRef] [Google Scholar] [Publisher Link]