AI SUPPORTED EMOTIONS ANALYSIS: A SYSTEM TO PROMOTE ENGAGEMENT IN DIGITAL LEARNING

Lucio Monterubbiano[1] and Chiara Moroni[2]
[1]*Skilla.com, Italy*
[2]*Amicucci formazione Srl - Skilla.com, Italy*

ABSTRACT

The paper presents a system of emotions and attention analysis during e-learning multimedia contents consumption.
The paper wants to explore the possible connections between facial expressions and user engagement while consuming multimedia e-learning courses such as tutorials, on-line quizzes and learning games. If a relevant connection emerges we will be able to better understand the attendant satisfaction and enhance the learning experience.

KEYWORDS

Artificial Intelligences, Computer Vision, Emotion Analysis, e-Learning Engagement, Facial Expression Recognition, Learning Profiles

1. INTRODUCTION

The project studies a possible appliance of artificial intelligence algorithms in the computer vision for the analysis of facial expressions during online e-learning sessions. A web technology based framework uses open sources libraries and open training models to implement a face recognition and expression recognition algorithms that process incoming video streams (from computer webcam) while the learners are attempting online courses via the same computer or device.

Those expressions are recorded several times in each seconds keeping records of the possible emotions and the attention level of the learner while consuming the content.

With the right amount of data we try to identify different behavior and the learning styles of the learner aiming to suggest the best content or make use of the most effective engagement techniques like for example a particular story telling or gamification strategy.

We identified 3 main reasons why this project can impact digital learning:

- By monitoring the level of attention and the emotive reactions in digital learning with a system with a relative low amount of bias from both teacher and learners we believe we can increase the worth of the user experience.
- If a relevant connection emerges between the user reactions and user appreciation we can have a measure of the effectiveness of engaging techniques (such as microlearning, storytelling, gamification, etc.) even without a direct feedback of the user.
- By monitoring the level of attention and the emotive reactions we hope to identify and categorize learning profiles or learning styles and build an effective recommendation system for adaptive learning dynamics.

2. DESCRIPTION OF THE PROJECT

Emotions play a fundamental role in life and also in learning. It is has been widely proven that positive emotions would endorse the learning process and would enhance the conceptual changes which are the bases. In the context of digital overload of our society in this era, the search for attention and focus is fundamental for the learning to be effective, especially in asynchronous e-learning sessions.

Due to this reasons we need strategies to monitor the attention level in order to create a better and more engaging content.

Thanks to the Google open source project called tensorflow we can use machine learning algorithms without creating ones from scratch. This platform has the peculiarity of evolving the precision and effectiveness of the task by consuming data. The algorithm would train itself analyzing real world inputs and would get better at doing its own task. This kind of algorithms aren't always the best choice because we already know, even before starting a project, that we will need to feed them with plenty of data to reach a good accuracy threshold. On the other hand in some field, like images or voice recognition, machine learning is able to reach a level of accuracy impossible for standard algorithms.

Those techniques are therefore used in our computer vision application for processing a video stream coming from the webcam of the user device. The video stream can be handled as a sequence of images, hopefully including details of the user face and expressions. Using another open source libraries called faceapi-js, we are able to detect the so called 68 points face model (figure 1).

The 68 points face model consists in identifying particular face landmarks in the every processed frame coming from the webcam video stream. Based on the work of Matthew Day (*Exploiting Facial Landmarks for Emotion Recognition in the Wild*) we can have a guess about the face expression by measuring some particular distances between the identified points (figure 1). Even if the face expression can be determined with a certain accuracy this doesn't mean that the user is actually experimenting that emotion with the same exactness. We are trying to understand if the correlation between the expressions we can detect with this tool and the appreciation of the user at the end of session can be meaningful.

It is not strictly important for our projects to guess the exact emotion in every frame, we are focusing more on the data set of emotion graphs and curves recorded during the online e-learning sessions (figure 2) in order to have an overview of the user experience during the attendance of a particular learning object.

In the end of the learning session we will propose the learner a quick multiple choice quiz and a 5 star evaluation assessment, in order to have 3 different variables that describes the same learning experience: the emotion graph, the quiz score and the five start evaluation.

With the multiple choice quiz we hope to have an indicator of the user understanding of the subject while with the 5 star evaluation we are asking the user her or his opinion about the learning object she or he has just viewed.

Comparing those 3 informations we are looking for a correlation, if there is any, between the emotion graph generated by the our tool and the other classical indicators we retrieved from the learners.

If the facial expressions analysis tend to be relevant, based on those collections and graphs, we can try to identify different learning outlines and learning styles. The assumption is that similar reactions to the same learning contents, drawn by emotions curves, can lead a similar learning attitude. Therefore, based on this kind of grouping and categorization, we will be able to propose and suggest the best contents or use engaging techniques to maximize the interest and the attention of the learners in the next e-learning sessions.

Note that the profile should be defined without an explicit question to the user, to minimize bias an prejudices, and it should be obtained from the person behavior in front of the pc or other device, which we hope to be more spontaneous.

Figure 1. The abstract 68 points face model and influential distances in the emotions detection model (*Matthew Day*)

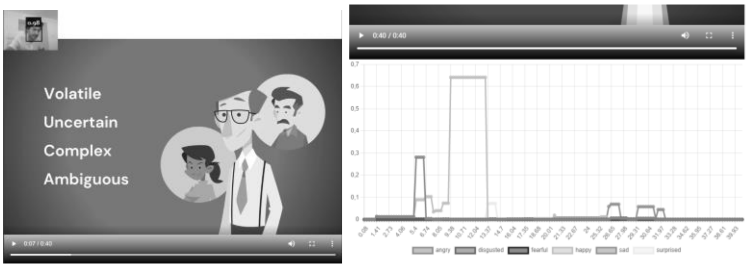

Figure 2. The 68 points face model detected on user with the guessed emotion and attention detection during the e-learning session

2.1 Open Questions

The literature of machine learning teach us that if we let the algorithm train itself with very few restrictions we can achieve unexpected results and explore game changing ideas. At the same time if don’t tune and constrain the training process with labels and rules we need and enormous amount of data to let the AI reach an acceptable performance.

For example, in our implementation of an adaptive learning system, should we label and tag the learning content with our “human” vision including our bias and prejudice or should we let the algorithm try a lot of different pattern and combinations in order to measure and discover the best recommendations?

Probably the first approach would be faster and better for early users, while the second can hopefully lead to better matches and associations in the long term.

3. CONCLUSION

In the continuous changing era, the continuous learning attitude it is a must in every sector and discipline.

Creating attractive learning content and optimizing the learning experience is a powerful weapon to keep learners engaged and involved.

We hope we can help measuring the effectiveness of the engagement with the help of AI, and then shaping and designing better the digital contents of the future.

REFERENCES

Amicucci, F. (2021). *Apprendere nell’Infosfera*, Franco Angeli, Milano

Ariely, D. (2009). *Predictably Irrational*, HarperCollins, New York

Behera, A., Matthew, P., Keidel, A., Vangorp, P., Fang, H., & Canning, S. (2020). *Associating Facial Expressions and Upper-Body Gestures with Learning Tasks for Enhancing Intelligent Tutoring Systems*. International Journal of Artificial Intelligence in Education (Springer Science & Business Media B.V.), 30(2), 236–270.

Huang, G. B., Ramesh, M., Berg, T., and Learned-Miller, E. (2007). *Labeled faces in the wild: A database for studying face recognition in unconstrained environments* Technical Report 07-49, University of Massachusetts, Amherst, Tech. Rep., 2007.

Kahneman, D. (2011). *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York

Lin, K. C., Huang, T.-C., Hung, J. C., Yen, N. Y., & Chen, S. J. (2013). *Facial emotion recognition towards affective computing-based learning.* Library Hi Tech, 31(2), 294–307.

Lucangeli, D. (2019). *Cinque lezioni leggere sull'emozione di apprendere*, Erickson, Trento

Matthew D. (2016). *Exploiting Facial Landmarks for Emotion Recognition in the Wild*, Available at: https://arxiv.org/pdf/1603.09129 (Accessed: 5 July 2022).

Shuo Y. (2016). Face Detection Benchmark. Available at: http://shuoyang1213.me/WIDERFACE (Accessed: 3 July 2022).

Vladimir Mandic (n-d). Face-api js library. Available at: https://www.npmjs.com/package/@vladmandic/face-api (Accessed: 11 July 2022).

Yee Phy Cho (n.d). Tensorflow face detection. Available at: https://github.com/yeephycho/tensorflow-face-detection (Accessed: 9 July 2022).