

EMOLEARN-"Las aventuras de Marco"- A Serious Game to train emotions to children with ASD*

Antonio Barba^{1,*}, Verónica Rufo¹, Giuseppe Iandolo¹, Esteban García-Cuesta²

¹Universidad Europea de Madrid

²Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, España

Abstract

This paper proposes the use of serious games for development of the emotional psychology of children with autism spectrum disorders. We exploit machine learning and Speech Emotion Recognition techniques to help children with ASD understand how to express their emotions at different day by day situations that they can be involved in. Marco's video game aims to stimulate metacognition and model adaptive behaviors, assertive communication, and self-regulation in daily social life. It is intended as a tool that a therapist can use to improve the learning process.

Keywords

Serious Games, Children with ASD, Speech Emotion Recognition, Machine Learning

1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by communication, and social interaction impairments, restrictive and repetitive behavioral patterns, interests, or activities with different levels of severity [13]. Child and adolescents with ASD show information processing, behavioral integration, and executive dysfunctions [14] with adaptive behavioral impairments in social contexts [15]. Environments overloaded with stimuli can lead to abnormal behavioral responses in children and adolescents with ASD and sensory modulation disorder, such as isolation, avoidance, or disruptive behaviors. Modeling sequences can entail an opportunity to understand the discomfort and develop social strategies to reduce it by asking for help from the environment in a socially appropriate way. Interactive educational games are designed to improve the social communication capabilities of children with ASD to detect and express emotions. EMOLEARN-"Las aventuras de Marco" aims to stimulate metacognition and model adaptive behaviors, assertive communication, and self-regulation in daily social life. In the first scenario, Marco learns to ask to play in the schoolyard without being intrusive, managing his anger at a possible refusal and asking the teacher for help. In the second, Marco learns to apologize to the people he may be bothering on a city bus. In the third scenario, Marco learns to communicate with the teacher when the light bothers him or there is much noise in the class. In the fourth, he learns to ask for a task in a class workgroup during his cast. Finally, in the fifth scenario, Marco learns to ask his classmates to talk to him individually

I Congreso Español de Videojuegos, December 1–2, 2022, Madrid, Spain

*Corresponding author.

✉ antonio.barba@universidadeuropea.es (A. Barba); veronicarufobaena@gmail.com (V. Rufo);

giuseppe.iandolo@psisemadrid.es (G. Iandolo); esteban.garcia@fi.upm.es (E. García-Cuesta)

🆔 0000-0002-5066-8637 (V. Rufo); 0000-0002-1215-3333 (E. García-Cuesta)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

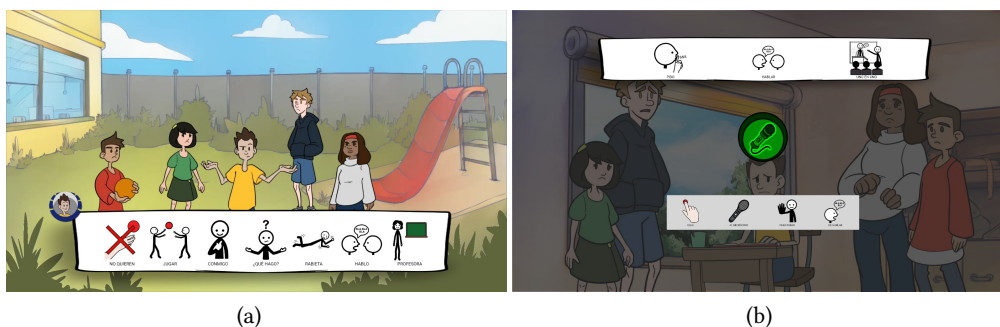


Figure 1: Screenshots of the game representing different situations that the student has to learn from. (a) Schoolyard scenario with pictograms that explains the current social context. (b) Marco explains to his roommates without anger that they should talk to him one by one.

and without overwhelming him. The first, second, and fourth scenarios aim to expose the player to a situational model of complex social interaction. The third and fifth scenarios aim to provide situational models that need to ask for help in case of overflow due to a sensory modulation disorder characterized by hypersensitivity to environmental stimuli.

2. Related Works

In recent years, much progress has been made to improve the communication skills for children with ASD thanks to machine learning emotion recognition techniques. This emotion recognition is based on both, facial gestures and voice. Thus, projects like Anwar and Milanova [1], LIFEisGAME [2] and FaceSay [3] present systems to identify facial expressions, faces and emotions both in expression and in detection for autistic children, while the Emotify project [4] recognizes the emotion through the voice in a talking educational game. In addition, there are projects that also take the social aspect into account. The ASC-Inclusion project [5] is a game experience that helps children with ASD to improve their socio-emotional communication skills, combining voice and facial and body gestures. Also, the Simoes et al. project [6] developed a serious game where the children became familiar with the routine of taking a bus. In our work we focus on the emotion of anger, since anger control as a manifestation of anxiety is a goal of any intervention in social skills in ASD and it is also one of the first emotions that children recognize. Moreover, by restricting the study to a simpler speech emotion recognition problem, the accuracy of the implemented model increases and, therefore, the chances of success in real settings.

3. EMOLEARN: Las aventuras de Marco

EMOLEARN-"Las aventuras de Marco" is a visual novel game where the main character is a 6-10 year old boy with ASD called Marco. The game consists of 5 everyday collaborative contexts associated with the 5 types of learning that are sought and kinematics are used to develop

them. At the end of the cinematic a decision process is proposed by means of pictograms with text (see figure 1). The child has to make a choice and if the option is correct then the process of oral exposition is opened where the child explains what happened. The video game evaluates the associated emotion and determines whether the child has spoken with or without anger. In the case of speaking in anger, the process has to be repeated until the child says it without getting angry. The Speech Emotion Recognition (SER) system includes spectral and prosodic features. The signal is framed into 20 ms windows to analyze its frequency content in a short time segment of a longer signal (this is a commonly chosen time window size but other sizes may be also a valid option). Following the authors recommendation at [10] the spectral features that we have used are: i) first 13 Mel Frequency Cepstral Coefficients (MFCCs) and their mean, standard deviation, kurtosis and skewness, the first and second derivatives of MFCC; ii) spectral centroid; iii) spectral flatness; iv) spectral contrast; and v) Linear Predictive Coding (LPC). The prosodic features represent those supra-segmental elements of oral expression which are elements that affect more than one phoneme and can't be segmented into smaller units, such as accent, tones, rhythm and intonation. The prosodic features that we use are the fundamental frequency (F0), intensity, and tempo. As result a total of 140 features are extracted using Praat [11] and librosa [12]. Different experiments were conducted using SVMs, Feed-Forward Deep Neural Networks (FFNN), and EXtreme Gradient Boosting (XGBOOST) and Cross-Validation (CV) methodology to obtain the most accurate model on anger vs. no-anger emotion classification task. The best model was obtained for SVM technique with an accuracy of 80%. Due to the difficulties autistic children have in managing attention, the game has been adapted to their needs according to the work of Frith's Central Coherence Theory [7]. This theory points out the difficulty that people with ASD have in analysing a set of stimuli as a whole, focusing their interest on details and increasing their capacity for fragmentation [9]. In order to improve children's attention, the following design decisions have been made for the visual paratranslation [8]: a simple and contrasting colour palette has been selected for the main character, the backgrounds have little information in order to focus the detail on the main action, a visual repetition of facial expressions and actions is sought to facilitate the relationship of situations with respect to the emotional attitude that occurs in the scenes, and the interface is direct without visual details that could distract from the main objective.

4. Conclusions and Future Works

This paper presents a fully functional serious games for development of the emotional psychology of children with ASD. The game is intended to be used by therapist in order to improve the emotional learning in social environments overloaded with stimuli. We expect that this will be a valuable tool in this context and it will be validated doing an A/B test experiment in a classroom of 20-30 children with ASD.

Acknowledgments

This work has been partially supported by European Commission Erasmus+ Strategic Partnerships for School Education (REF.KA201-063086).

References

- [1] Anwar et al. 2016. Real Time Face Expression Recognition of Children with Autism. International Academy of Engineering and Medical Research (2016).
- [2] Abirached, B. et al. 2011. Improving communication skills of children with ASD through interaction with virtual characters. IEEE 1st SeGAH (2011).
- [3] Hopkins, I. et al. 2011. Avatar assistant: improving social skills in students with an ASD through a computer-based intervention. *Journal of autism and developmental disorders* 41, 11 (2011), 1543–1555.
- [4] Rouhi, A. et al. 2019. Emotify: Emotional Game for Children with Autism Spectrum Disorder based-on Machine Learning. In 24th International Conference on Intelligent User Interfaces (IUI '19 Companion), March 17–20, 2019, Marina del Rey, CA, USA. ACM, New York, NY, USA, 2 pages.
- [5] Schuller, B. et al. The state of play of ASC-Inclusion: An Integrated Internet-Based Environment for Social Inclusion of Children with Autism Spectrum Conditions. In Proceedings 2nd International Workshop on Digital Games for Empowerment and Inclusion (IDGEI 2014).
- [6] Simões M, Bernardes M, Barros F, Castelo-Branco M. Virtual Travel Training for Autism Spectrum Disorder: Proof-of-Concept Interventional Study. *JMIR Serious Games*. 2018 Mar 20;6(1):e5. doi: 10.2196/games.8428. PMID: 29559425; PMCID: PMC5883078.
- [7] Frith, U. (1989). *Autism: Explaining the enigma*. Blackwell.
- [8] Méndez González, R. and Calvo-Ferrer, J. R. (2017). *Videojuegos y [para]traducción: aproximación a la práctica localizadora* (Editorial Comares).
- [9] Shah, A. and Frith, U. (1983). An islet of ability in autistic children: A research note. *Journal of Child Psychology and Psychiatry*, 24, 613-620.
- [10] Schuller, B., Wöllmer, M., Eyben, F., and Rigoll, G. (2009) *Spectral or Voice Quality? Feature Type Relevance for the Discrimination of Emotion Pairs*. The Role of Prosody in Affective Speech (S. Hancil, ed.), vol. 97 of *Linguistic Insights, Studies in Language and Communication*, pp. 285–307, Peter Lang Publishing Group, 2009
- [11] Jadoul, Y., Thompson, B., and de Boer, B. (2018). *Introducing Parselmouth: A Python interface to Praat*. *Journal of Phonetics*, 71, 1-15. <https://doi.org/10.1016/j.wocn.2018.07.001>
- [12] McFee, Brian, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. *librosa: Audio and music signal analysis in python*. In Proceedings of the 14th python in science conference, pp. 18-25. 2015.
- [13] American Psychiatric Association - APA. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. American Psychiatric Pub.
- [14] Murcia, C. L., Gulden, F., and Herrup, K. (2005). *A question of balance: a proposal for new mouse models of autism*. *International Journal of Developmental Neuroscience*, 23(2-3), 265-275.
- [15] Baron-Cohen, S. et al. (1999). *Social intelligence in the normal and autistic brain: an fMRI study*. *European journal of neuroscience*, 11(6), 1891-1898.