

Augmented Reality based social skills intervention system for ASD children using Video Self-Modelling

Sapna Jain¹ M Afshar Alam²

^aDepartment of Computer Science, Jamia Hamdard, New Delhi

^aDepartment of Computer Science, Jamia Hamdard, New Delhi

Abstract: Autism Spectrum Disorder (ASD) is a lifelong, intricate condition that hinders social and communication abilities and affects around one in hundred children under the age of 10 in India. Tailored training based on individual needs may not always be feasible for individuals with ASD. Children with ASD struggle with social interaction and empathy. The goal is to help individuals with autism improve their learning and communication skills, which is a dream for their families. There is evidence indicating that children with ASD show a preference for images of themselves over others, highlighting the necessity for an Augmented Reality-based intervention system to improve social skills. The current ASD systems in India do not include an AR environment, personalized social narratives, or Automatic Video self-modelling (VSM). The AR systems currently in place only cover basic social greeting skills and do not focus on addressing mind blindness and socio-communication skills in individuals with ASD. Additionally, existing ASD with Gesture recognition systems are not reliable as the actions performed by children with autism are unpredictable. The current ASD systems in India lack an AR environment, personalized social narratives, and Automatic Video self-modelling (VSM). The AR systems currently implemented only focus on basic social greeting skills and do not address mind blindness and socio-communication skills in individuals with ASD. Furthermore, existing ASD with Gesture recognition systems are unreliable due to the unpredictable actions of children with autism. The primary aim of the proposed model is to develop an efficient Augmented Reality solution for evidence-based intervention in Autism Spectrum Disorder (ASD) to assist children (Mild to Moderate level) in overcoming mind blindness. This will involve training in various socio-communication skills, creating cartoon animation videos with injected self-images during social narrative sessions, and integrating Eye gaze detection. This will allow individuals to learn by observing themselves in cartoon videos while performing specific tasks. The eye gaze of the learner is detected during live social narrative session to avoid the deviation from the social narrative and alert will be given. This proposed model will perform robust multimodal Gesture detection, Voice detection and Automatic Video Self-Modelling in practice session to help Autism Spectrum Disorder (ASD) children to practice and review their activities. The proposed Automatic Video self-modelling (VSM) system will summarize the practice video of the learner by identifying the timestamp at which successful gesture and voice is detected. This investigation will be carried out with Narrative story content creation for various social communication skills and support for baseline phase and intervention phase for experimentation and validation from National Institute for Empowerment of Persons with Multiple Disabilities (NIEPMD) thereby providing better learning and communication to benefit the Autism children.

Keywords: Augmented Reality, Personalization, Cartoon animation, Autism Spectrum Disorder, Automatic Video Self Modelling, Eye gaze detection, Gesture Recognition

1. Introduction

The paper addresses the hassle of cognitive blindness or mental incapacity in youngsters with autism spectrum ailment (ASD) with verbal and nonverbal verbal exchange skills. Mental blindness is described as the absence of emotional reciprocity. The first-rate way to assist teenagers with ASD examine social competencies is to use the web to enhance their skills. The aim of this mission is to increase a reality-based approach to assist youngsters with ASD (mild to moderate) collect primary conversation capabilities thru an online, evidence-based method. Avatars show the social conversation abilities made viable via ultra-

modern structures the usage of augmented fact and self-representation. because summary questioning is a location that is frequently tough for humans with ASD, they might also have subject making use of newly received know-how in one-of-a-kind approaches in a giant setting. It is additionally frequent for VR structures to download simulations. While growing avatars can be difficult, many VR structures are totally black or have avatars that do not seem to be plenty like the user.

2. Literature Review

The Multimodal Adaptive Social Interaction in Virtual Environment (MASI-VR) VR intervention machine was once developed with the aid of the Looxid lab in South Korea for young people with ASD. By combining eye gaze, EEG signals, and peripheral psychophysiological signals, the gadget makes use of a multimodal VR-based social interplay platform to pick out the kid's emotional kingdom and cognitive patterns. In order to overcome the fears that can forestall humans with autism from residing their lives to the fullest, professionals from Newcastle University and Third Eye Neuro Tech laboured collectively to create The Blue Room [11], a customized 360-degree environment. Through a partnership with Yale University's School of Medicine, the Centre for Brain Health and the Child Study Centre purpose to assist younger adults with ASD reap social and monetary independence via utilising digital truth technology. Managing conditions such as job interviews, neighbour disputes, and even relationship used to be made less complicated by using it. Children with autism can now examine in productive surroundings thanks to the efforts of Chandra Reka Ramachandiran and Nazeen Jomhari from the University of Malaya and Taylor's University shore campus. This was once finished via the use of photograph change verbal exchange machine (PECS) technological know-how to accumulate statistics and in-person interviews to create a digital surroundings model, which was once because of this assessed thru quantitative evaluations. There is education for the events tasks. An app for digital actuality that teaches children how to pass the road safely was once created via Israel's University of Haifa. To assist schools, therapists, parents, and help body of workers assist young people with ASD, Floreo[7] presents a digital fact gaining knowledge of and social competencies application. Through social memories and an interactive digital sensory environment, college students who use the app research necessary social capabilities like reciprocity and anticipatory planning. They additionally analyse how to independently control their emotions. Dr. Patrick Bordnick, dean and professor at Tulane School of Social Works, developed the VR-qualisest vita app for children and adults with autism. In order to assist customers, increase their social skills, verbal exchange skills, and comprehension of interactions in a range of contexts, which includes the domestic and school, this software immerses customers in a sensible environment. According to the critiques of job counsellors, Rizzo, a lookup professor at the University of Southern California in Los Angeles, created a digital application for job interviews that appreciably elevated the interviewing competencies of individuals who had autism. Using a digital lecture room with eight avatar spectators, Davis developed a comparable application to assist youngsters with autism exercise public speaking. If the speaker does no longer preserve eye contact with the avatars, they will begin to disappear, encouraging humans to appear round the room alternatively of simply forward. Virtual fact (VR) video games are designed to mimic competencies like sharing and properly communication, giving possibilities for both. Giuliana Guazzaroni and Anitha S.Pillai collaborated to create a digital fact software for college students with autism spectrum disorder. Wearing a specific head-mounted device, college students can connect, interact, locate 360-degree academic resources, and entire tutorial assignments in these digital surroundings that focuses on instructing social skills. With help from the education services at Action for Autism AFA, a pinnacle autism middle in the nation, Pratibha Vellanki has been developing low-priced computer-based treatments or video modelling for humans with autism in India via adapting applications like TOBY for the nearby populace. While creating a Hindi model. The teens used each language choices to advance their skills; the Hindi mannequin produced marginally higher results. In partnership with the Department of Science and Technology and the Anvay Research and Trainings Center, the Action for Autism AFA has initiated a mission to enhance things to do and content material for computer-assisted early intervention packages that are culturally appropriate. This new fabric can be delivered to already-released software, such as the multimedia framework e-Saadhya. Children with ASD had their behavioural issues and developmental capabilities efficiently addressed by means of the ComDEALL trust. ComDEALL is founded on enhancing language, motor, and daily dwelling skills. They prefer educating autistic youth thru computer-based therapy.

3. Proposed Work

The main objective of this project is to create an effective Augmented Reality solution for evidence-based intervention in ASD that helps children with ASD improve their social communication skills and overcome mind blindness. This technology utilizes a Kinect sensor to incorporate self-portraits into animated videos with social stories in order to practice different communication abilities. This project is made up of two parts: a Social narrative session and a Practice session. Social narrative is a portrayal of social behaviour scenarios for children with ASD, with the learner being the main character in an animated cartoon video. Evidence suggests that children diagnosed with ASD show a preference for images of themselves compared to others. The skilled instructor at NIEPMD, alongside the Co-PI, will write the story narrations for different Social Communication skills outlined in Table 1. This is an intervention for ASD that is based on evidence, where one learns by observing themselves on video while performing a specific behaviour. The user's eye location is a helpful indicator of attention. This system predicts the user's focus on the screen, which is a valuable signal of attention. The calculation relies on where the eye-gaze vectors estimated from each eye intersect. If it is found that the eye-gaze is not focused on the monitor, the image will be illuminated in red in different colour; blue not included. After grasping the storyline of a specific situation, the learner will engage in a practice session where their movements are recorded by the Kinect camera. The live footage is displayed on the system screen known as the viewing window, and an animated agent appears next to it to provide feedback and support to the learner. During this session, a strong multimodal system is used to recognize the gestures and speech of the learner. After the learner performs the gestures for a specific scenario, successful gestures and voice are detected through Deep Neural Network. Video snippets are stitched together using automatic Video Self Modelling to create a video for the autism child to review their actions. This is known as feed forward VSM, enabling the learner to enhance a specific desired social behaviour.

Table 1: Social Communication skills

SKILLS IN COMMUNICATING
Skills in communication through conversation
1. Start a conversation at the proper time.
2. Begins discussing a particular subject
3. Inquiries about information using "Wh" interrogative words.
4. Answers questions beginning with "Wh"
5. Provides a range of remarks on the subject matter while discussing.
6. Brings individuals together for the first time.
7. Introduces people to each other
8. Concludes discussions in a suitable manner
Compliments
1. Provides suitable compliments to colleagues.
2. Receives compliments in a suitable manner.
3. Requests a favor with proper etiquette
4. Apologizes on their own accord.

The system consists of two parts i.e. Social narrative and Practice session. A social narrative is a lively tale that depicts the main character interacting with various cartoon characters in a specific setting, including meeting, greeting, complimenting, and apologizing. This section contains a narrative explanation that learners must use to develop their social communication skills by observing the scenario. Evidence-based methods are intervention strategies that have been proven effective through research. Focused intervention practices aim to target one specific goal for individuals with ASD. One example of a practice grounded in evidence is employing social narratives for behaviour intervention. Social narratives are brief stories that clearly explain social scenarios and give instances of how to react appropriately within a particular culture. These stories are personalized to the student and are narrated in the first person. Social narratives interventions aim to assist learners in comprehending pertinent social cues and promoting suitable social responses within their surroundings. This social narrative system enhances a written story with animated visuals showcasing the social skill. Additionally, the learner is portrayed as the central figure in the narrative, utilizing

animation to showcase social skills in different social scenarios with adults and peers. "A social story is a brief tale that explains a situation by focusing on important social cues and typical reactions, giving students precise and detailed information about what happens in a situation, and the reasons behind it." A social narrative is created for each individual child, containing elements that are important to them and capture their interest. The story targets to decorate the person's comprehension, furnish comfort, and provide appropriate responses for the particular situation. Prior to growing a social story, we ought to think about sure aspects:

a) The kid's ability degree

b) Figuring out the precise talent or scenario to tackle upon. The device being counselled teaches intervention for a set of social skills. If a toddler with ASD is delivered to someone, the educating narrative for guiding him in fabulous behaviour in that state of affairs will be developed.

i) Narration for in the back of "Introduces himself or herself to anybody new" I would possibly meet anybody new. They may say "hi" to me. I can additionally say "hi" to him or her. I can begin the dialog by means of pronouncing "Hi" to him or her first if he would not say "Hi" to me first. Then, he or she would possibly inquire, "How are you?" I ought to reply with "Fine" and inquire about his health. I may additionally have to ask him first if he does not ask. When I introduce myself to the person, I may additionally additionally shake their hand.

ii) Comparatively, to "end a discussion", the story would be: "They may additionally say, "Bye, best to meet you," or I can say, "Bye, satisfactory to meet you," shake their hand, and give up the conversation,"

iii) Portrayal for the commendation skills beneath the agenda: "Assuming I see any person did his/her task properly and returns subsequent to ending it, I cost his work by using providing reward "great job". In the match that he did not come and let me understand after fruits of undertaking, I, at the stop of the day, may additionally begin and discover his work via supplying reward s "generally excellent". After that, he may renowned the praise and say, "Thank you." If he does no longer reply to the compliment, I will ask him to say "Thank you" in response. I can also likewise say 'welcome' to him after I get tons appreciated.

The clarification for the ability "Apologizes": "I may additionally have a look at anybody doing a task." He would possibly comprehend that something went incorrect with his work and say, "Sorry for the mistake." I may want to say, "That's fine. I would possibly begin by using saying, "You are wrong," if he did not apologize. He responds, "Sorry for the mistake." I should respond, "That's fine. "The usual narrative session is depicted in Fig. 1 and it is similarly personalized in the following way:

Table 1: Social Narrative	
STEPS	PROCESS
Step-1	Narrative text displayed on screen.
Step-2	Personalization → Eye gaze detection steps
Step-3	Pre-Processing
Step-4	Eye region detection
Step-5	Feature Extraction
Step-6	Feature detection and tracking using Supervised Descent method

Personalization: The proposed system would produce the animated cartoon narrative to illustrate the scenario, and the autism children instructor at NIEPMD would provide the text narratives of a particular scenario (such as an introduction to a person) in accordance with their method of training the children. The system incorporates self-images by individualizing and substituting the participant's image for the face of the main character. In our method, we colorize the body and replace the head of the main character to portray the learner. The Unity 3D Game Engine software was used to create this animated cartoon narrative. Models, textures, and animations are used to visually represent characters in 3D games. Mecanim, a system in Unity, makes it simple to create animations, blend them together, and retarget humanoid animations between models. The work process in Mecanim can be parted into three significant stages.

1. Import and preparation of assets. This is finished by, with outsider apparatuses, like Max or Maya. The Mecanim features are not relevant to this step.

2. There are two methods for setting up a Mecanim character: Humanoid character arrangement. Mecanim has a

unique work process for humanoid models, with broadened GUI support and retargeting. The arrangement includes making and setting up a Symbol and tweaking Muscle definitions. o Nonexclusive person arrangement. This applies to things like creatures, animated props, animals with four legs, and so on. Although retargeting is not possible in this case, you can still make use of Mecanim's extensive feature set, which includes everything that is discussed below.

3. Rejuvenating characters. This involves exposing Animation Parameters, establishing State Machines and Blend Trees, establishing interactions between animation clips, and controlling animations from code. The story for the "Hi" scenario could look like this as shown in fig. 2 below:



Figure 2: Narrative for 'Hi' [1]

This device indicates how to speak socially with adults and friends in a range of situations. The learner and different characters are animated in such a way that they exhibit how to wave, speak, or have a dialog or supply a compliment. A range of scenarios like adults greeting children, kids greeting adults, and adults greeting adults—were chosen due to the fact they are comparable to conditions that are often encountered in day-to-day life.

Session I: Observation of the gaze:

A treasured mark of consideration is the place of the client's eye stare. The Kinect digicam returns range pix of the expert and the consumer all through a meeting. Intraface analyses these photos in actual time. The place of every eye's student and the gaze course are estimated by way of Intraface. We estimate the user's gaze role on the reveal the use of this data. The main goal is to instruct two cascaded regressors (CR) to realize the scholar core the usage of the Supervised Descent Method (SDM) and Random Cascaded-Regression Copse (R-CR-C). The first CR, which is acknowledged as the Face Cascaded Regressor, F-CR, is skilled to apprehend a set of 32 landmarks that correspond to facial points, whilst the 2d CR, which is regarded as the Eye Cascaded Regressor, E-CR, is educated to understand the internal and outer eye corners as nicely as the scholar facilities (six landmarks). The gauge relies upon on the crossing factor of the beams fashioned by using the eye-stare vector gauges from the two eyes. The photo is illuminated purple if it is decided that the eye gaze center of attention is now not on the monitor. Rather, in the match that the client's nonetheless up in the air to be on the PC, the photo is lit blue. Eye gaze detection is solely used at some point of the narrative and now not at some stage in the exercise session when the person is shut to the computer. This is due to the fact our intention is to hold the reader's interest during the narrative.

Session II: Practice session

The social narrative is observed with the aid of the exercise session. It offers the pupil a risk to put what they've discovered in the preceding social narrative into practice. Particularly, the motive of the exercise session is to foster the initiating of social interplay and the terrific response to it. An interplay agent and a viewing window make up the exercise session. The Kinect camera's real-time video feed is displayed in the viewing window. The consumer can consequently study themselves carrying out the action. Because this maintains the learner related to the story, the interplay sellers are person or peer characters that the learner has already encountered in the narrative. In order, the animated characters appear. The session requires the learner to reply to the characters' social conversation capabilities in order to cross on to the subsequent story. For instance, the learner has the alternative of initiating a greeting or responding to a character's greeting. Text prompts provide instruction and

encouragement. The consumer is rewarded with shiny sounds and lights for making profitable wave gestures. Since the exercise session can be replayed the usage of one of the buttons at the pinnacle of the display screen and the average glide is shown in Fig. 3 the instructor controls the pace of the session, giving the student as many chances as they need.

Table 2 : Practice Session	
STEPS	PROCESS
Step-1	. Video input from user
Step-2	Gesture Detection
Step-3	Voice Detection
Step-4	If voice detection responds to Interaction Agent, then initiate conversation.
Step-5	If conversation initiation process is success, then set parameters of onset, offset locations, brightness, color and light, then do segmentation.
Step-6	If step-5 is successful then proceed to next story
Step-7	If step-5 is unsuccessful then provide another chance.

Acknowledgements: The authors acknowledge Fund for Improvement of S&T Infrastructure (FIST) LAB and Promotion of University Research and Scientific Excellence (PURSE) LAB is one of the Flagship Infrastructure programs of Department of for providing the infrastructure to work on the proposed methodology.

4.Conclusion: The proposed machine will allow the autism children to have a greater reading and verbal alternate in the actual world which appears to be a dream for their dad and mother and family. The have an impact on to this work will seem to be for benefits for affected communities. This mission will make contributions positively to the autism humans and grant a rapid appreciation and gaining know-how of the social verbal trade Genius the use of augmented reality. Autism kids can examine shortly with the aid of staring at oneself on video performing a target behaviour. The mission of this undertaking is to make the autism kids to analyse and communicate easily with the society and additionally assist them no longer to make social error due to Mind blindness that cause all types of problems. The transport of the self-modelling makes it less complicated for the ASD kids to undertake and for environment friendly and fast learning. Thus, the individual's perception will be increased and make him extra cozy and maybe advocate identical terrific responses for the state of affairs in question.

References

- [1] N. Uzuegbunam, W. Wong, S. Cheung, and L. Ruble, "MEBook: Multimedia social greetings intervention for children with autism spectrum disorders," in IEEE Transactions on learning technologies, Oct 2016.
- [2] C.R.Ramachandran, N.Jomhari, S.Thiyagaraja, M.Maria, "Virtual reality based behavioural learning for autistic children," in Electronic Journal of e-Learning vol.13 2015.
- [3] E. Bekele, J. Wade, D. Bian, J. Fan, Amy Swanson, Z. Warren, N. Sarkar, "Multimodal adptive social intervention in virtual environment (MASI-VR) for children with autism spectrum disorders(ASD)," in IEEE Int. Conf. Virtual reality, March 2016.
- [4] N.Josman, P.L.Weiss, M.Ben-Chaim, H.Friedrich, "Effectiveness of Virtual reality for teaching street crossing to children and adolescents with autism," in Int.journal for disability development and education, Jan 2008.
- [5] N.Gelbar, T.Buggey, "Video self-modeling as an intervention strategy for autism spectrum disorders," in psychology in schools, Jan 2012.
- [6] A. L. Wainer and B. R. Ingersoll, "The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders," Res. Autism Spectrum. Disorders, vol. 5,2011.
- [7] K.Kaur and Seadeba Pany, "Computer-Based intervention for autism spectrum disorder children and their social skills," in scholarly research journal for humanity sciences, Sep 2017.
- [8] K.Oyedotun and A.Khashman, "Deep learning in vision-based static hand gesture recognition," in Conf. Neural computing and applications, April 2016.
- [9] Guillaume Devineau ; Fabien Moutarde ; Wang Xi ; Jie Yang,"Deep Learning for hand gesture Recognition on skeletal data," in IEEE Int. conf on automatic afce and gesture recognition, May 20. Xuemei Ren ; Huichao Li," Gesture recognition method based on deep learning," in IEEE youth academic annual conference of Chinese assoc., july 2018.
- [10] Pratibha Vellanki," Applied machine learning for personalised early intervention in autism," May 2016.

- [11] Morag Maskey," Virtual reality therapy treats autism phobias" Newcastle University, Feb 2019
- [12] Sri Harsha Dumpala, Sunil Kumar Kopparapu, "Improved Speaker Recognition System for Stressed Speech using Deep Neural Networks", IEEE 2017
- [13] Guangming Zhu, Liang Zhang, Peiyi Shen and Juan Song," Multimodal Gesture Recognition Using 3D Convolution and Convolutional LSTM", IEEE Access 2017.