Second Delivery Advanced HCI

Giuseppe Grisolia, Michele Minniti

Final idea: Studio Assistant

- Multimodal studio assistant to help with AAC students with special needs
- Image generation through speech to visualize the subjects of the explanation
- Gaze control and emotion recognition to assess difficulties or give positive feedback with gamification setup
- Enhancing teaching-learning experience and overcome educational barriers

Improvements from last presentation

- Slide numbers!
- Clearer ideas about the software and functionalities
- Enhanced some aspects about our idea

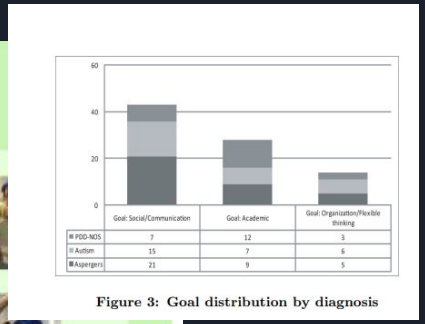


Figure 3: Goal distribution by diagnosis

g. 3 MyDentist:


RESEARCH Open Access

The impact of gamification on students' learning, engagement and behavior based on their personality traits

Rodrigo Smiderle[3], Sandro José Rigo[3], Leonardo B. Marques[1], Jorge Arthur Peçanha de Miranda Coelho[2] and Patrícia A. Jaques[3*]

*Correspondence: patricia.jaques@gmail.com
[3]Programa de Pós-graduação em Computação Aplicada (PPGCA) - University of Vale do Rio dos Sinos (UNISINOS), São Leopoldo, Brazil
Full list of author information is available at the end of the article

Abstract
The gamification of education can enhance levels of students' engagement similar to what games can do, to improve their particular skills and optimize their learning. On the other hand, scientific studies have shown adverse outcomes based on the user's preferences. The link among the user's characteristics, executed actions, and the game elements is still an open question. Aiming to find some insights for this issue, we have investigated the effects of gamification on students' learning, behavior, and engagement based on their personality traits in a web-based programming learning


Neuropsychiatric Dise

The role of au
communicatic
current status and future tr

This article was published in the follow
Neuropsychiatric Disease and Treatme
19 September 2016
Number of times this article has been

Teresa Iacono[1]
David Trembath[2]
Shane Erickson[3]

[1]Living with Disability Research Centre, La Trobe University, Bendigo, VIC, Australia; [2]Menzies Health Institute Queensland, Griffith University, Gold Coast, QLD, Australia; [3]Living with Disability Research Centre, La Trobe University,

Background: Augmentative
for children with autism, often
minimally verbal. Our aim was
(up to 21 years), and then consic
based autism interventions targ
Design: We completed a syst
ERIC) as well as forward citat
intervention efficacy research
March 2016 in peer-reviewed

0:52 ✓0

PROFUMA PUZZA

Methodology to identify goals and needs

We started by analyzing several papers on the topic of the idea proposed and in particular regarding "ASD child learning and their needs".

We analyzed the existing applications and technologies that support these kind of users and statistics about the learning process.

what resulted from the analysis of the papers is:

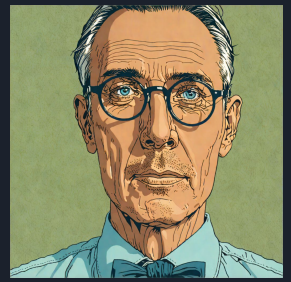
- make products more portable
- Most existing products are only in the educational sector, and do not take into account the emotional and social sphere
- in a survey reported in one of the papers, aimed at parents of ASD children twenty-two respondents (19%) specifically requested that software be designed with fun in mind = they felt that developers should consider creating learning experiences like games. (GAMIFICATION)

Personas

*Generated with Stable Diffusion

CHILD USER DESCRIPTION:

- **Name:** Rosie Ortiz
- **Age:** 7 y. o.
- **Characteristics:** Rosie is a young child with Autism Spectrum Disorder (ASD) who struggles with communication in her day-to-day activities. Rosie loves colorful visuals, somewhat repetitive patterns, and simple interfaces. She is sensitive to loud sounds, so she prefers quiet, calm settings.
- **Goals:** Learning at school in a comprehensible and fun way in order to not get bored.
- **Challenges:** Needs positive reinforcement to not get frustrated nor distracted. Gets frustrated if she fails to achieve her goal after a few trials.
- **Key stakeholders**: Parents, Teachers

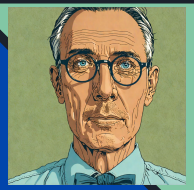


*Generated with Stable Diffusion

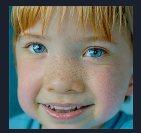
TEACHER USER DESCRIPTION:

- **Name:** Marc Reds
- **Age:** 65 y. o.
- **characteristics:** Marc is a support teacher specialized in ASD children. Has more than 30 years of experience in the field and has found interesting experimenting with new technologies to enhance his teaching methods.
- **Goals:** Support the child with ASD integrating innovative technological tools to improve communication and learning and monitor his progress.
- **Tech familiarity:** Simple but effective, knows how to run a software on his pc or an app for tablets.
- **Challenges:** Needs a comprehensive interface, variety of possibilities to improve communication with the child

Detailed functionalities → “Interactive Explanation”



“albero”



What it does:

- Converts teacher's spoken instructions or explanations into vivid, illustrative images in real-time.

Purpose:

- Enhances comprehension by providing visual support aligned with verbal instructions.
- Bridges the gap between abstract concepts and understanding through visualization.

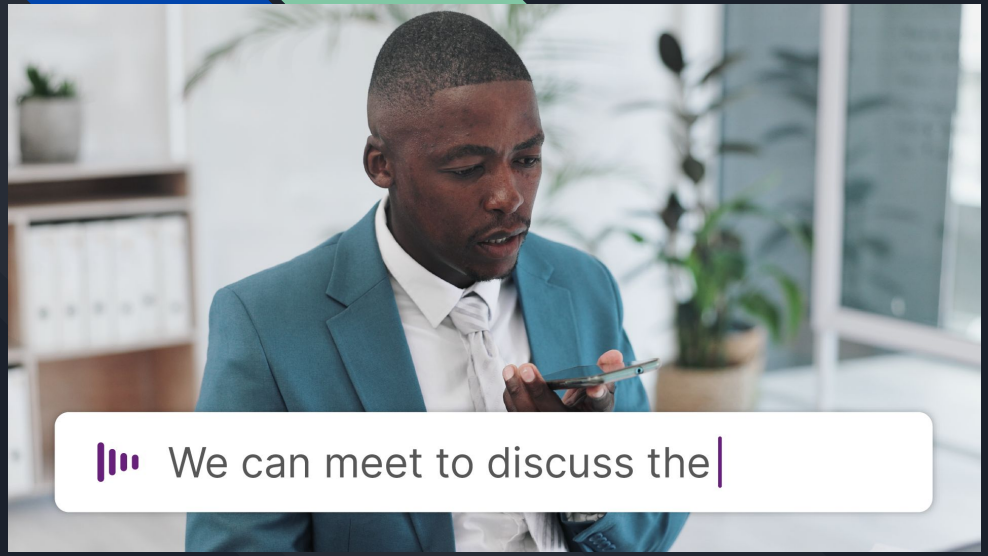
How it works:

- Speech recognition identifies keywords and contexts from the teacher’s speech.
- AI generates images dynamically based on detected inputs.

Example:

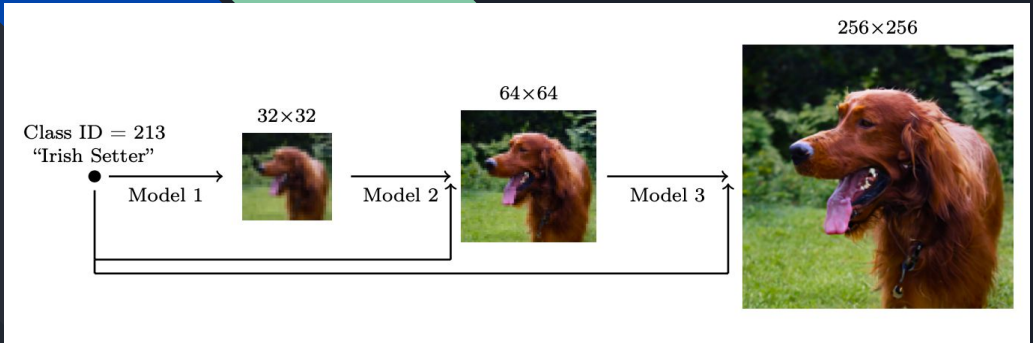
- Teacher says: *"Imagine a forest with tall trees and a clear blue sky."*
  - The software instantly generates an image of a serene forest.

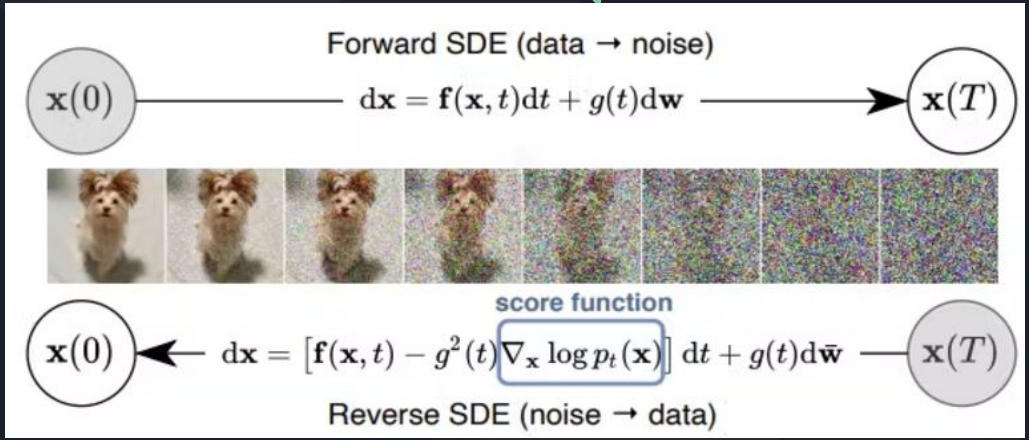
Whisper Tiny



- Whisper Tiny is an **ultra-lightweight speech-to-text (STT) model** designed for high efficiency and real-time transcription.
- Compact
- Multilingual support
- Real time performance
- Robust to noise
- Minimal resource requirements
- Capable of working in pipeline with stable diffusion

Stable Diffusion





- Diffusion Models (DM) show great image quality and diversity in the generation.
- Using a pre-trained autoencoder Latent DM achieves great results with resolutions till 1024x1024.
- Stable Diffusion: Based on a latent diffusion model operating in a lower-dimensional latent space for computational efficiency.
- The model asks about 4GB of VRAM but eventually we can use quantization to improve efficiency

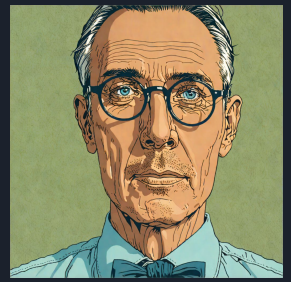
Personas

*Generated with Stable Diffusion

CHILD USER DESCRIPTION:

- **Name:** Rosie Ortiz
- **Age:** 7 y. o.
- **Characteristics:** Rosie is a young child with Autism Spectrum Disorder (ASD) who struggles with communication in her day-to-day activities. Rosie loves colorful visuals, somewhat repetitive patterns, and simple interfaces. She is sensitive to loud sounds, so she prefers quiet, calm settings.
- **Goals:** Learning at school in a simple and fun way in order to don't lose attention.
- **Challenges:** Needs positive reinforcement to don't get frustrated nor distracted. For example: she gets frustrated if she fails to achieve her goal or perform exercise after a few trials.
- **Key stakeholders**: Parents, Teachers

*Generated with Stable Diffusion



TEACHER USER DESCRIPTION:

- **Name:** Marc Reds
- **Age:** 65 y. o.
- **characteristics:** Marc is a support teacher specialized in ASD children. Has more than 30 years of experience in the field and has found interesting experimenting with new technologies to enhance his teaching methods.
- **Goals:** Support the child with ASD integrating innovative technological tools to improve communication and learning and monitor his progress.
- **Tech familiarity:** Simple but effective, knows how to run a software on his pc or an app for tablets.
- **Challenges:** Needs a comprehensive interface, variety of possibilities to improve communication with the child

based on Rosie needs…

“EMOTION RECOGNITION”

…to control her mood

What it does:

- Detects emotional states in real-time through facial expressions.

Purpose:

- Helps teachers adapt their explanations based on the child’s emotional responses.
- Early identification of struggles or disengagement.

How it works:

- Uses AI models trained on diverse emotional datasets.
- Provides real-time feedback to the teacher on the child’s emotional state.

Key Benefits:

- Promotes emotional awareness.
- Promotes empathetic teaching approaches.

based on Rosie needs...

“DISTRACTION RECOGNITION”

...to control that she doesn't lose attention

What it does:

- Tracks the child’s gaze to monitor focus and engagement.

Purpose:

- Alerts teachers if the child is losing attention or focus.

How it works:

- Analyzes gaze direction and duration using a webcam

Applications:

- Helps redirect focus through interactive activities.
- Teacher can adapt lesson content dynamically to REgain attention.

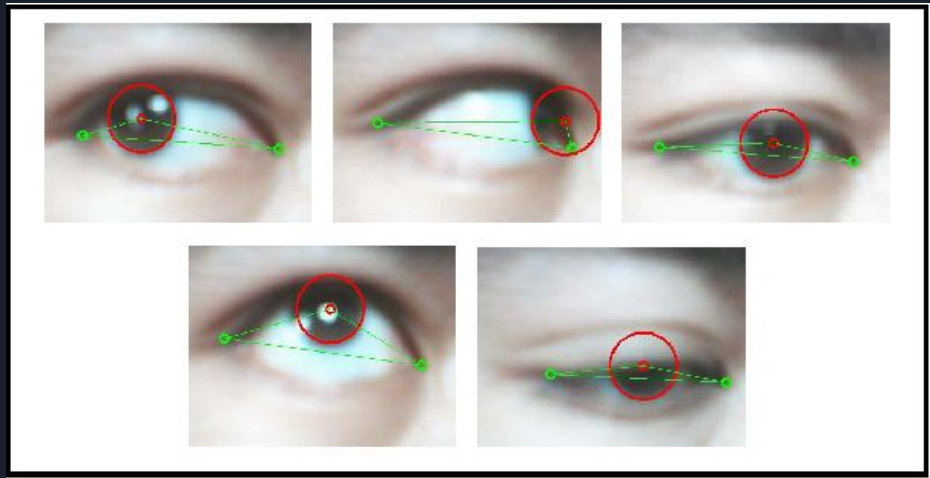


Figure 12. Results of the first experimental

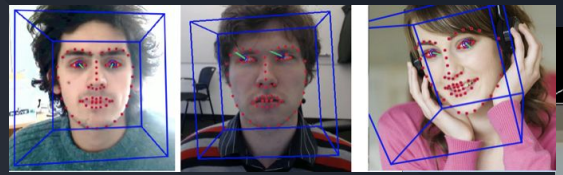
Synergy of Emotion and Gaze Recognition

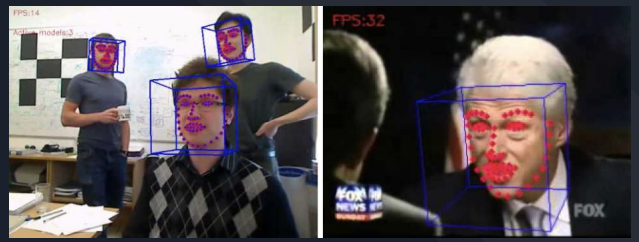
Combined Functionality:

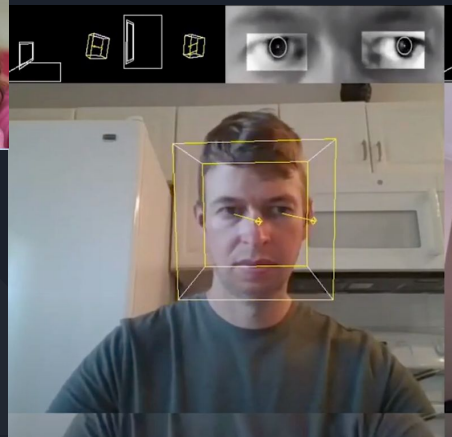
- Emotion and gaze data are combined for a better understanding of child's engagement.
- Example:
  - Gaze indicates focus, but emotion shows frustration → may be potential struggle with content.
  - Gaze and emotion indicate disengagement → prompts re-engagement strategies.

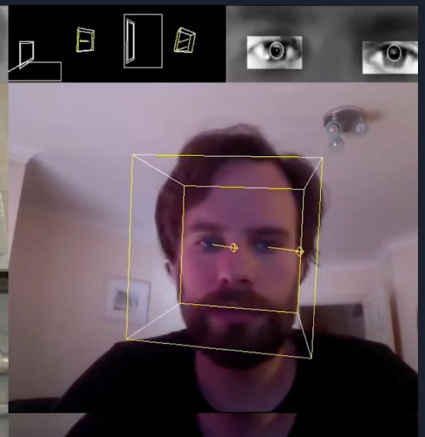
Outcome:

- Provides important feedback to teachers for adjusting lessons.
- Improves the overall learning experience.









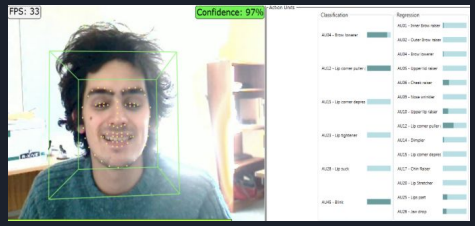
Openface

OpenFace is a state-of-the art tool intended for automatic facial behavior analysis and understanding.

The main functionalities are:

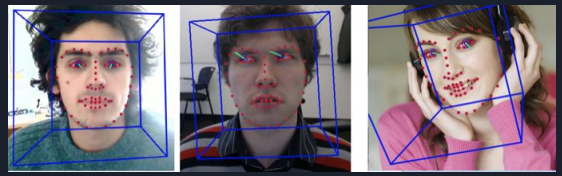
- **FACIAL LANDMARKS DETECTION**:

Identifies key points on the human face (such as eyes, nose, mouth…ecc) using computer vision and machine learning techniques. These points are useful for our idea because we can analyzes facial expression for emotion recognition functionality



- **FACIAL AU RECOGNITION**:

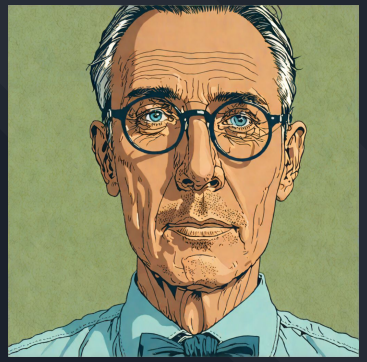
Analyzes and recognizes facial action units (AUs), which represent specific facial muscle movements. This system is useful to study expressions and micro-expression in applications such as psychology, medicine, entertainment and also useful to our scope
= Emotion and attention recognition functionality.



- **EYE-GAZE ESTIMATION**:

Determines the direction of a person’s gaze by analyzing eye position and corneal reflections. Also useful to our scope = Emotion and attention recognition functionality.

Scenario
TEACHER USAGE



CONTEX:
Marc Reds, support teacher, uses the app to explain a school concept to a child with ASD, making the lesson more interactive thanks to the “interactive explanation” function. He can also monitor the emotion and concentration of the child with the emotion/distraction recognition.

STEPS:

1. Marc opens the app on the Rosie’s profile;
2. Marc select the “Interactive Explanation” mode;
3. Marc activates the “Distraction & Emotion recognition” mode;
4. Marc start the explanation. Example: “Today we learn what trees are. Trees have branches, leaves and roots!”;
5. The app generate in real-time image of a tree, branches, leaves and roots;
6. The app also provide a real-time feedback about the concentration and emotion of Rosie.
7. Marc consult the concentration/emotion feedback of rosie and continues explanation and talks about other topics, which are turned into pictures for Rosie;
8. Marc assign an exercise to Rosie about the topic explained before.
9. … MOVING TO SCENARIO OF STUDENT USAGE …

Scenario

STUDENT USAGE

CONTEX:
Rosie uses the app to complete an elementary exercise, given by the teacher, while the app monitors his emotional state to detect any difficulties or distractions.

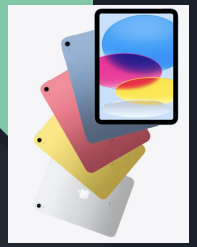
STEPS:

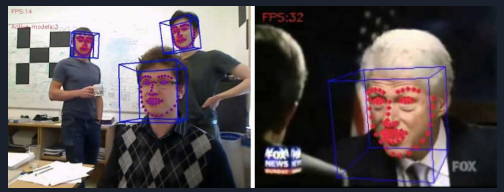
1. Rosie receives the exercise from teacher;
2. The app presents an instruction that explains what to do, and images to better understand the task;
3. Rosie start to complete the task
4. The functionality “Distraction & Emotion recognition” previously activated by teacher is running and is analyzing the gaze of rosie for distraction recognition and the facial expression to monitor her emotion and feelings during the exercise;
5. If Rosie gets DISTRACTED, the app emits a signal, like a colored pop-up appear. For example: a cartoon character telling her to continue the exercise.
6. Same stuff in case Rosie becomes SAD/NERVOUS/AGITATED the app emits a signal like a colored pop-up appear. For example: a cartoon character telling her to relax, maybe take a break, and retry.
7. In the end, the app provides the outcome of the exercise (right/wrong) and the report on the distraction and the emotions felt by Rosie → FOR THE TEACHER

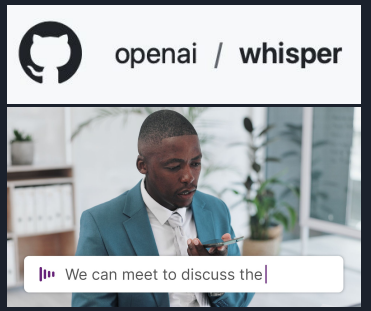
Hardware and software tools

TO DEVELOP OUR IDEA OUR IDEA

which basically will be an application







AS HARDWARE TOOLS:

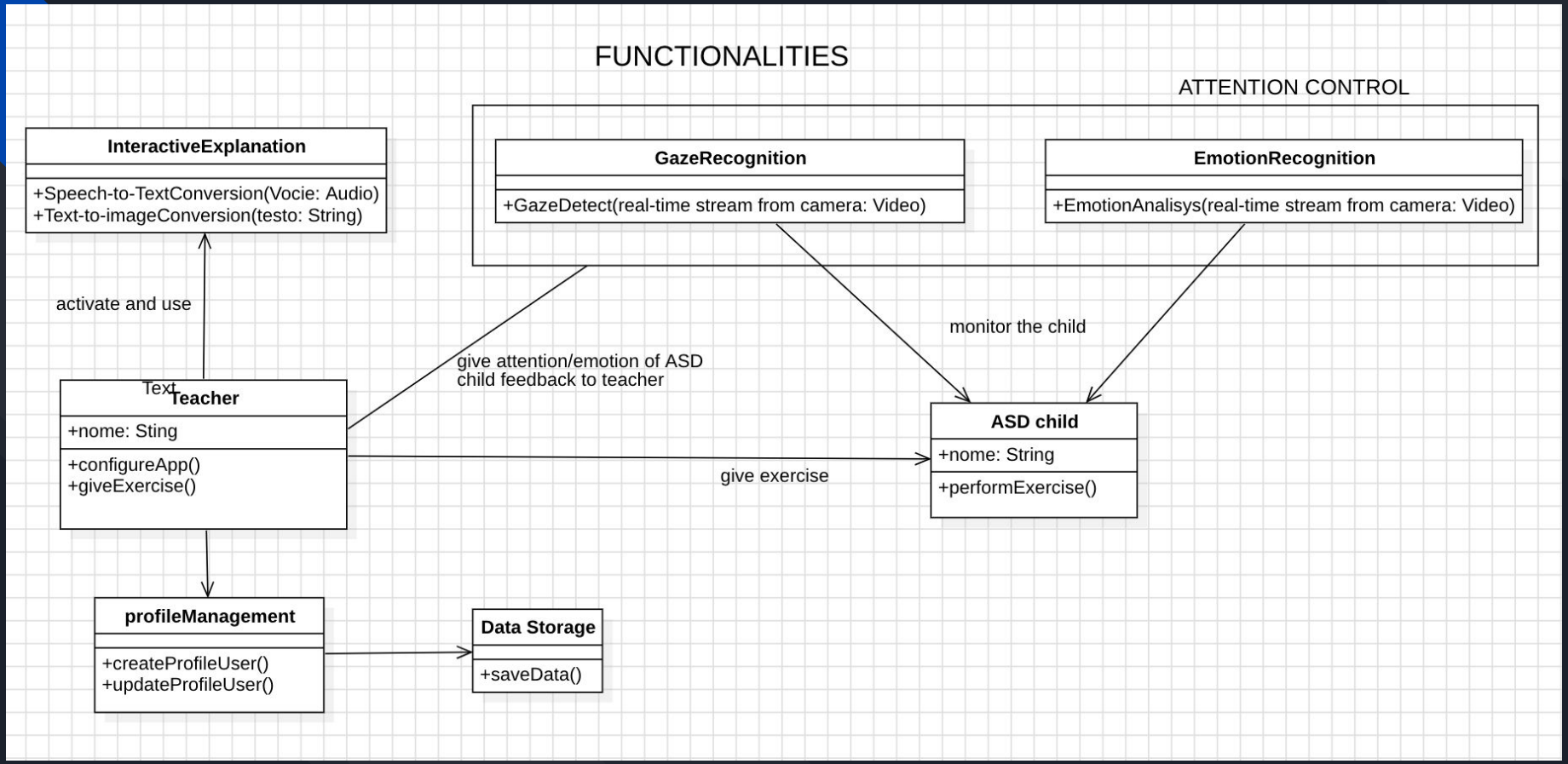
1. LAPTOP
2. TABLET

(with a VRAM of at least 4GB)

AS SOFTWARE TOOLS:

1. OpenFace tool-kit
2. OpenAI whisper tiny version
3. Stable Diffusion for image generation

UML MODEL


FUNCTIONALITIES
ATTENTION CONTROL
InteractiveExplanation
+Speech-to-TextConversion(Vocie: Audio)
+Text-to-imageConversion(testo: String)
GazeRecognition
+GazeDetect(real-time stream from camera: Video)
EmotionRecognition
+EmotionAnalisys(real-time stream from camera: Video)
activate and use
Text
Teacher
+nome: Sting
+configureApp()
+giveExercise()
give attention/emotion of ASD
child feedback to teacher
monitor the child
ASD child
+nome: String
+performExercise()
give exercise
profileManagement
+createProfileUser()
+updateProfileUser()
Data Storage
+saveData()

References

1. *Akamsha Timande, Pallavi Borse, Vaishnavi Lande, & A.G.Sharma. (2024). Speech to Image Generation by Stable Diffusion Model. SSGM Journal of Science and Engineering, 2(1), 89–91. Retrieved from https://ssgmjournal.in/index.php/ssgm/article/view/116*
2. *Boluk, Nursena ; Kose, Hatice. / **Evaluating Gaze Detection for Children with Autism Using the ChildPlay-R Dataset**. 2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition, FG 2024. Institute of Electrical and Electronics Engineers Inc., 2024. (2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition, FG 2024).*
3. *Paweł Tarnowski, Marcin Kołodziej, Andrzej Majkowski, Remigiusz J. Rak, Emotion recognition using facial expressions, Procedia Computer Science Volume 108, 2017. Retrieved from https://www.sciencedirect.com/science/article/pii/S1877050917305264*
4. Iacono, T., Trembath, D., & Erickson, S. (2016). The role of augmentative and alternative communication for children with autism: current status and future trends. *Neuropsychiatric disease and treatment*, 2349-2361.
5. Putnam, C., & Chong, L. (2008, October). Software and technologies designed for people with autism: what do users want?. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility* (pp. 3-10).
6. Khaleel, F. L., Ashaari, N. S., & Wook, T. S. M. T. (2020). The impact of gamification on students learning engagement. *International Journal of Electrical and Computer Engineering*, *10*(5), 4965-4972.
7. Bondioli, M., Buzzi, M. C., Buzzi, M., Chessa, S., Jaccheri, L., Senette, C., & Pelagatti, S. (2024). Guidelines for research and design of software for children with ASD in e-health. *Universal Access in the Information Society*, *23*(4), 1909-1930.
8. Baltrušaitis, T., Robinson, P., & Morency, L. P. (2016, March). Openface: an open source facial behavior analysis toolkit. In *2016 IEEE winter conference on applications of computer vision (WACV)* (pp. 1-10). IEEE.
9. Sauer, A., Lorenz, D., Blattmann, A., & Rombach, R. (2025). Adversarial diffusion distillation. In *European Conference on Computer Vision* (pp. 87-103). Springer, Cham.
10. Cakmak, M. C., Okeke, O., Spann, B., & Agarwal, N. (2023, May). Adopting parallel processing for rapid generation of transcripts in multimedia-rich online information environment. In *2023 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (pp. 832-837). IEEE.