A conversation supporting technology

WIME – What Is My Emotion?

A conversation supporting technology for people with social-emotional agnosia

By Thale Kirkhorn, Svenja Lys Forstner and Kaja Hustad Bendiksvoll

Introduction

In the project, we focus on a support solution for a condition called "social-emotional agnosia", also known as "social blindness" (mdmedicine, 2011). It describes the inability to interpret facial expressions, body language, and tone of voice. A condition common amongst people on the autism spectrum. Individuals with this condition experience difficulties with accurately understanding another person's emotions in social situations. Picking up non-verbal communication during conversations is especially challenging. A device such as WIME could assist in social situations and improve the flow of the conversation.

> Autistic children often […] find it hard to use emotion to understand social interactions. They might not notice when others are upset or angry. They might show less concern for others and have less ability to comfort others or share emotions (*Emotional development in autistic children*, 2020).

Over the course of the semester, we have worked on a project exploring a way to support people with social-emotional agnosia in conversations. The idea was to use facial recognition software and a camera positioned at approximately eye height. This way the camera has the same view of the other person as the user, and there is less chance of something disrupting the view of the camera. The findings of the software should be displayed as visual feedback on a watch. This gives the user the ability to see what emotion the other person is expressing.

Implementation and technical details

The initial plan was to implement the project using a Raspberry Pi and a respective camera module to connect the camera, facial recognition software, and smartwatch. After some work in setting up the Raspberry Pi to stream video from the camera in real-time, we chose to use a laptop and web camera instead in order to connect everything. The tools used for the final result are OpenCV (Open Source Computer Vision Library), TensorFlow, and the Keras API (Dutta, 2021). The emotion recognition system uses a pre-trained machine learning model to recognize facial expressions from a real-time video stream. The software detects the emotion of one person at a time. At the moment, six different emotions are recognized: happy, angry, sad, surprise, fear, and neutral. When an emotion is detected a python script creates a file

with JSON data of the current emotion. This file updates each time a new emotion is recognized. This JSON file is hosted on a local webserver to be accessible via the local network. The smartwatch is then connected to the website via Bluetooth and retrieves the emotion from the JSON data.

For the visual feedback, we use the programmable smartwatch “Bangle.js” by Espruino. In addition to the touch display, the watch also has sound and vibration. Additionally, it can detect the user's pulse (Williams & Ikkala, 2021). For each of the emotions, we created a different screen output. Six different emojis represent one emotion each. A text is also displayed underneath, to further clarify the emotion detected. Because of the simple output, the person wearing the watch can easily read the display with a quick look. The watch updates every few seconds, so whenever a new emotion is detected, it is shown on the watch immediately. One might think that it could be distracting for the other person to see their emotion displayed on a screen throughout the conversation. However, the screen is made in such a way that it looks black when you look at it from the side.

Fig 1: Visual output on watch when the emotion sad is detected.

Possible issues and privacy concerns

Naturally, there are some issues that came up during development which are going to be discussed in the following. We also wanted to take the possible privacy concerns into consideration as this topic is especially important for the user. First, the decision to use a video stream was made to ensure the privacy of the users by not recording and saving user

data. The local web server creates a log of every emotion detected, however, it is not possible to connect it to the corresponding person. Furthermore, the log is cleared after each session. This avoids any further privacy concerns, as we may be dealing with sensitive information without people's permission.

Another possible issue we've discussed is that people may express emotion differently depending on the situation. Will the software, for example, identify sarcasm? It is especially vital if the individual struggles to understand sarcasm, and it will not assist the situation if it is not operating correctly. We have also looked into whether emotion AI can have any racial prejudice, or the possibility of lower accuracy if facial attributes deviate from what the network is trained on. The technology should be capable of deciphering facial expressions, assessing speech patterns, tracking eye movements, and evaluating neurological immersion levels, which will provide a more accurate result. It is something we wish to improve in the future.

Further development plans

In the future, it would be helpful to incorporate ways to use body language and voice intonation for the visual feedback on the watch because it is likely to generate a more precise output if combined. Also, improving facial recognition software to recognize and interpret a wider variety of emotions, e.g. embarrassment, envy, or guilt, could help significantly improve the quality of the conversation. So far, the software can only identify six different basic emotions. Still, in a real-life conversation, there are many nuances to a person’s emotions. The last thing would be to add an alternative feedback solution that would be suitable for blind or visually impaired people as visual feedback would not be sufficient. We further wish to keep testing and doing more research on the focus group. Doing interviews through an autism organization, for instance, can give us more insight into the field of consumers. We believe this feedback is essential. We did not get the chance to do user testing because of time limitations, but we presented the project to experts in the field and got some insightful feedback and thought-provoking questions. Even if this was not perhaps the group that would use WIME daily, it is still beneficial to get an insight from them. For further development, we wish to have more detailed and structured interviews with a specific focus group that would use a product such as WIME in the future.

In conclusion, the basic idea of the technology has been implemented, and it is waiting to be tested in different environments and situations to be judged about its actual usability. It is important to emphasize that WIME will not save any data from the consumers, in order to protect their privacy. There are several ways to improve the project's current state, and room for further development.

Literature

Dutta, J. (2021, 06 22). *Raspberry Pi Based Emotion Recognition using OpenCV, TensorFlow, and Keras*. Retrieved from https://circuitdigest.com/microcontroller-projects/raspberry-pi-based-emotion-recognition-using-opencv-tensorflow-and-keras

mdmedicine. (2011, 04 14). *Agnosia*. Retrieved from https://mdmedicine.wordpress.com/2011/04/14/agnosia/

raisingchildren.net.au. (2020, 12 07). *Emotional development in autistic children*. Retrieved from https://raisingchildren.net.au/autism/development/social-emotional-development/emotional-development-asd

Williams, G., & Ikkala, T. (2021, 10 04). *Bangle.js - Espruino*. Retrieved from https://www.espruino.com/Bangle.js