NOVO NORDISK: DIGITAL COMPANION FOR THE CHRONICALLY ILL

Students: Sajjad Atabi, Valerie Chan, Ishan Dane, Zongjin Li, Mark Long, Evan O'Neill, Jin White

Problem Statement

- Studies have estimated that approximately ⅓ of people who have been diagnosed with a serious, life-changing chronic illness or disease will experience symptoms of depression as a result of their illness [1]. Furthermore, such illnesses have seen an increase in recent years, and as such an increasing number of people have experienced these symptoms. Compounding these symptoms, the isolation caused by these illnesses can result in serious negative effects on the patient's health.

Requirements

- The objective of this project is to develop a standstill digital companion device that can provide information, support, and companionship to bring positive impact to chronic disease patient. This is done through primarily focusing on mental, physical, and emotional support.
- The device should be able to encompass a wide variety of patients and conditions and provide personalized feedback based on each patient's unique needs.
- Requirement 1: Accurate speech to text conversion which generates the input for the companion and text to speech for output.
- Requirement 2: Robust NLP algorithm that can accurately and coherently respond to a given input
  - The NLP model should be able to accurately engage in dialogue with the user, which includes high content relevance as well as contextual consistency.
- Requirement 3: Comprehensive Machine learning platform for personalized responses
  - Should remember trends in the users behaviors and even recall past topics to make the user feel that it actively listens/cares.
- Requirement 4: Approachable and friendly design
  - The housing for the companion should be designed to put the patient at ease
    - Friendly, Non-abrasive

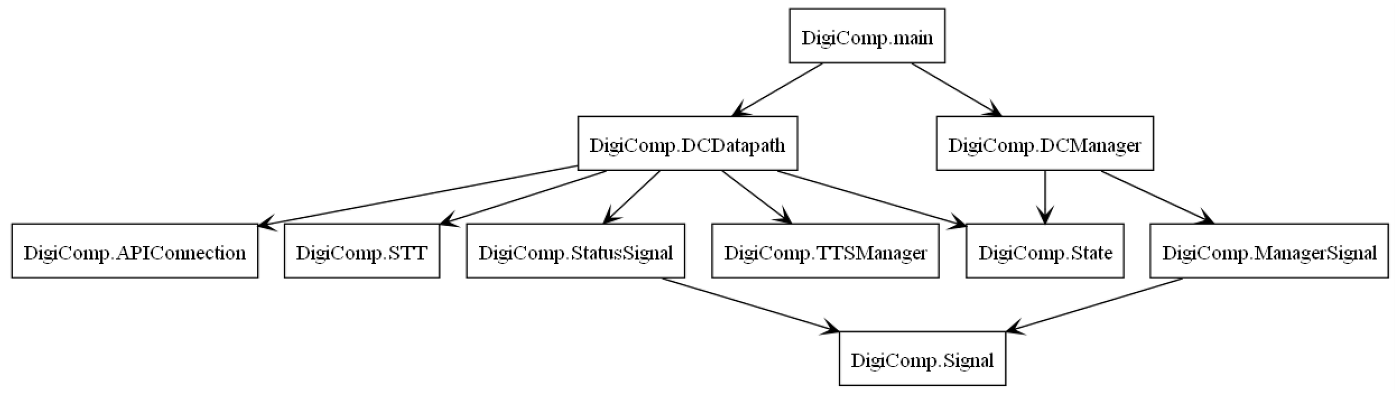

Figure 1. Block Diagram of the Software Architecture

Implementation

Speech To Text and Text To Speech

- The Speech-to-text Input is a phrase from the user which is transformed into text for the companion to parse with an NLP algorithm. Input speech needs post-processing before being passed off to the .txt file and NLP algorithm to correct inaccurate words given the context.
- Text-to-Speech takes the output from the API's, or any text input from an external location, and generates a audio file which is used to respond to the user. This function utilizes the TacoTron2 TTS Model and once initialized can generate and output a realistic voice response in less than 3 seconds.

NLP Companion REST API

- Our REST API provides an abstraction between the user and two other API's:
  - OpenAI's GPT3.5-turbo for NLP and human-like responses.
  - IBM Watson Emotion API to label companion responses with an emotion used to determine the eye-emotion to show.
- Two devices have authorization to access the API, the Robot and the iOS App. The iOS App is capable of toggling control to either take control of the API or release control to the Robot. This way, only one device can communicate with the companion at a time.

iOS Application

- The iOS application is a supplementary feature aiming to provide a portable version of our hardware-based digital companion. The software platform digital companion shares the same API kernel with the hardware platform version, thereby achieving consistency and integrity in information interaction.

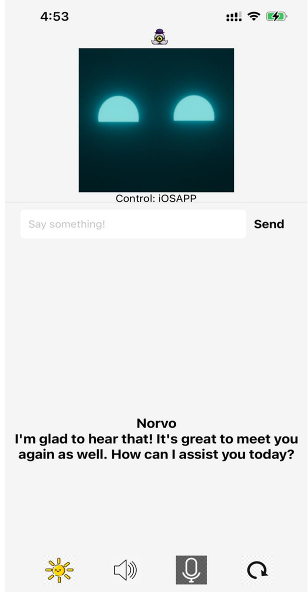

Figure 2. UI of the Digital Companion App

Computer Vision

- The companion uses a MediaPipe Pose Detection model along with OpenCV2 video capture and processing to allow it to identify if a human is in frame. It then follows their movements to maintain 'eye-contact' and mimic the head movements of a real person in conversation.

Hardware Design

- The overall physical components of the companion includes a LCD screen, speaker, capacitive sensors, and several motors along with a 3D printed housing to create an interactive experience for users.
- Eye contact and movement play a vital role in social and emotional communication as seen in clinical studies [2]. We placed special emphasis on using animated eyes to primarily convey emotions and enhance understanding and listening. By implementing animated eyes, we aim to create a more engaging and immersive experience for users, evoking a wide range of emotions and ultimately improving the overall communication process.

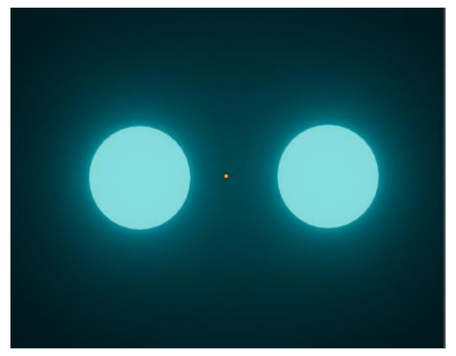

Figure 3. Screenshot of the animated eyes

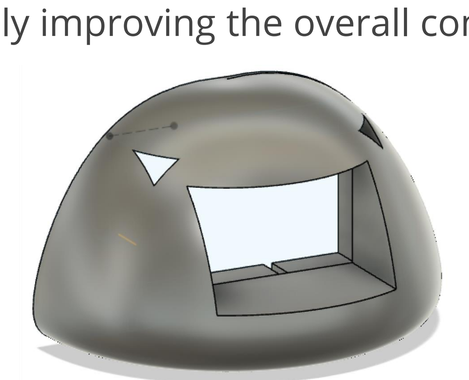

Figure 4. CAD model of the head of the companion housing

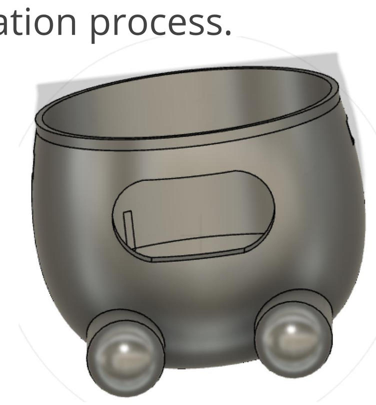

Figure 5. CAD model of the body of the companion housing

Discussion / Future Work

- Further improvements for the digital companion:
  - Improve chassis to use a more durable and aesthetically pleasing material
  - Expansion of TTS to include multiple different voices
  - Improvement of SST algorithm to increase accuracy and speed
- Further usability and user testing are also crucial in order to continue to develop the digital companion:
  - Identify emotions through animated eyes
  - Physical design of the digital companion
  - Overall user experience while interacting with the physical digital companion and the app

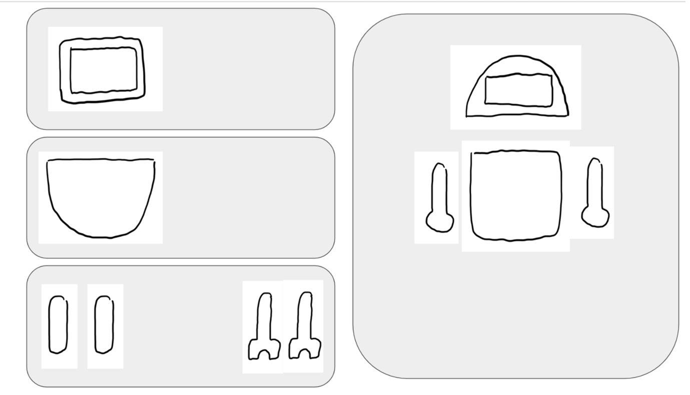

Figure 6. Screenshot of the interactive activity for physical design during the user testing

Conclusion / Results

- Natural Language Understanding (NLU): The digital companion was successfully able to interpret and understand natural language input, enabling seamless and intuitive communication with users.
- Emotion Expression via Screen: The companion's screen dynamically displays various emotions, allowing it to express a wide range of emotional responses and creating a more personalized and engaging user experience and when paired with the computer vision platform allows the companion to "follow" a user as they move around a space .
- Speech Processing: By using the TacoTron2 model for TTS the companion is successfully able to generate an audio response to a user's prompt within a few seconds. We tested another TTS model called TortoiseTTS which provided a even better voice generation but took significantly longer

Figure 7. Final prototype of the Digital Companion

Scan the QR to see a video demo of the Digital Companion!

QR code

Acknowledgement

Special thanks to John Canevari and Karolynn from Novo Nordisk, and Prof. Kim Ingraham from the University of Washington for all the guidance, support, and feedback for this capstone project.

References

[1] "Coping with chronic illnesses and depression," WebMD, https://www.webmd.com/depression/guide/chronic-illnesses-depression (accessed May 22, 2023).
[2] [1] M. S. U. E. Jodi Schulz, "Eye contact: An introduction to its role in communication," MSU Extension, https://www.canr.msu.edu/news/eye_contact_an_introduction_to_its_role_in_communication (accessed May 22, 2023).

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

FACULTY ADVISERS: Professor Kim Ingraham, Mingfei Chen

INDUSTRY MENTOR: John Cavenari

SPONSOR: Novo Nordisk