An Empathetic Conversational Humanoid Robot for Elementary School Children

Can commercial humanoid robots, infused with artificial intelligence (AI), recognize emotions in children based on tonalities of speech? Previous work in human-robot interaction (HRI) has shown robots trained and configured with the right datasets and models, can detect emotion in people based on the tonalities of speech [1]. However, humanoid robots capable of expressing empathy for children, through auditory conversation, is another challenge to explore. To implement empathy, the three layers of empathy must be considered-cognitive, emotional, and compassionate [2] (Fig 1). From volunteering at Prices Fork Elementary, I began to see the gaps of education that students are experiencing during COVID-19, and from my harsh circumstances at home, I want to develop an approach that can be used for STEM education and to support mental health of children.

Aim 1: Explore cognitive empathy between children and humanoid robots.

Aim 2: Have a humanoid robot express emotional empathy by classifying children's emotions through a trained audio classifier using deep learning principles.

Aim 3: Conduct an in-person case study, where the humanoid robot is communicating with children virtually (online) and "in-person" to see how well the robot acts from their emotions.

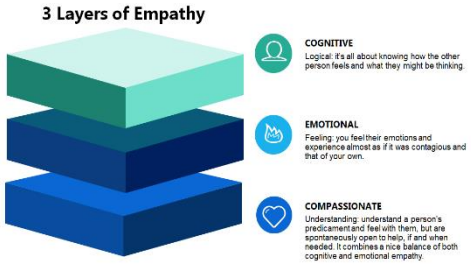

Figure 1: Showcasing the 3 layers of empathy: cognitive, emotional, and compassionate [2]

Experimental Setup: Softbank's Pepper robot has been commercially available since 2014, and is equipped with a multimodal means of communication, and a free, open source development environment. The TRIM-Lab has a Pepper robot allowing me to conduct the proposed research within a three-year timespan. **The goal is to create a conversational robot capable of understanding and recognizing speech to converse with children, which in turn, will also help further innovation and inspiration of AI**. The following activities are proposed for individual aims:

Aim 1: Students at Prices Fork Elementary are currently being taught in a hybrid setting, during COVID-19, where they are taught both in person and online all using Google Chromebooks. During the Fall of 2020, I have branched resources from the TRIM-Lab to the elementary school, to showcase educational technology the TRIM-Lab has to offer, such as Pepper. With the new robotic curriculum I have developed, and with the STEM director's permission, I will implement the first aim of my study to learn from students over their perception of Pepper, and measure cognitive empathy between children and Pepper in December 2020. From the Robot Theatre project, we plan to first showcase Pepper to students in the classroom from three one-minute videos of Pepper performing various scenes to students. Then, the elementary school students will virtually interact with Pepper via live stream in class on their Google Chromebooks. We have already submitted for approval of the study to Virginia Tech's Institutional Review Board (IRB). I am certified to conduct such research from the Collaborative Institutional Training Initiative. After students interact with the Pepper robot, they will anonymously be asked a quantitative survey over their perception of Pepper to understand how they felt about the robot from the movie clips and see how well Pepper addresses their emotions.

Aim 2: To program Pepper to recognize emotions from speech, one would have to fit, train, and test a large auditory dataset into Pepper. The dataset which will be used for the study is

EmoReact, an emotion-based dataset of 1102 videos of children from Youtube Reaction Channels, who are between the ages of four and fourteen years old [3]. It is annotated with 17 emotions, including six basic emotions (happiness, sadness, surprise, fear, disgust, and anger), which will be the main emotions focused to develop to fit, test, and train on Pepper. I want to utilize EmoReact to work on an optimized audio classifier that learns directly from just tonalities in audio recordings. This requires a deep learning (DL) model, to fit, train, and test the dataset on, which has been completed with the Ryerson Audio-Visual Dataset of Emotional Speech and Song (RAVDESS) but not for EmoReact [4]. I will implement this deep learning model with EmoReact to have Pepper recognize emotions based on tonalities of speech. After configuring the dataset with the model, I will transcribe movements and iterations Pepper will perform for the classroom, when recognizing different emotions.

Aim 3: Pepper will be used inside of the elementary school classroom and be livestreamed as a case study to see how well Pepper can express compassionate empathy for children. In the classroom, Pepper will classify their responses, as cognitive empathy by guessing their emotion of how they feel. Then express understanding of the student's emotions in the classroom, by classifying their emotion through conversation, showcasing emotional empathy. Finally, Pepper will show displays of compassionate empathy through gestures of visual and sounds after acknowledging and understanding children's emotions.

Publications: One of the outcomes of the proposed research will be publications:

1. 2021 October Ro-Man Conference (deadline 03/14/21).
2. 2022 Human Robotics Interaction Conference (deadline 10/21).
3. 2023 ACM Computer Human Computer Interaction (CHI) Conference (deadline 09/22).

Intellectual Merit: While in graduate school, I will study child-robot interaction (CRI). Data collected from children is rare and implementing a humanoid robot for children which can understand emotions in children is an avenue to explore. With the given resources of open access software, present resources at (University), and a feasible plan within a three-year timespan, this research will produce beneficial outcomes and create an inspiring impact for younger students. The research will put a framework for establishing two-way empathy between robots and children, by addressing the 3 stages of empathy for better understanding and communication, and a platform for other researchers to work.

Broader Impact: Commercial robots such as Pepper are already applied in industries such as retail, hospitality, and entertainment. However, the future avenue of humanoid robots which can genuinely express empathy cognitively, emotionally, and compassionately have wide applications. Not only could it benefit and enhance STEM education, but also mental health in children, nursing, and general healthcare. This research will pave the future for developing students for STEM careers and make AI more approachable, therefore, inspiring future leaders.

References:

[1] Lim, Angelica & Okuno, Hiroshi. (2014). The MEI Robot: Towards Using Motherese to Develop Multimodal Emotional Intelligence. Autonomous Mental Development, IEEE

[2] Powell, P. A., & Roberts, J. (2017). Situational determinants of cognitive, affective, and compassionate empathy in naturalistic digital interactions. *Computers in Human Behavior*, *68*, 137–148.

[3] B. Nagarajan and V. Ramana Murthy Oruganti, "Cross-Domain Transfer Learning for Complex Emotion Recognition," *2019 IEEE Region 10 Symposium (TENSYMP)*, Kolkata, India, 2019, pp. 649-653

[4] M. G. de Pinto, M. Polignano, P. Lops and G. Semeraro, "Emotions Understanding Model from Spoken Language using Deep Neural Networks and Mel-Frequency Cepstral Coefficients," *2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, Bari, Italy, 2020, pp. 1-5