

Non-verbal communication empathy assessment using extended OMG-Empathy dataset

Leandro H. de S. Silva^{1,2}, Letícia Portela¹, Nathália R. C. da Silva¹, Agostinho A. F. Júnior¹, Bruno J. T. Fernandes¹

¹ University of Pernambuco, Brazil

² Federal Institute of Paraíba, Brazil

Empathy is the ability to understand, perceive, and respond appropriately to the emotional state of someone else. Empathy is one of the most important aspects of human communication, and an essential asset for social agents. In human-robot interaction, empathy assessment can lead to better decision-making since this interaction can be more natural and human-like. However, one of the problems of automatic empathy modeling is that an agent cannot rely on perception of emotional expressions alone. When analyzing only text, the context of the situation matters because the same sequence of words can have different interpretations depending on the context. In this study, we will explore the use of Skip-RNNs to allow the analysis of long-term context on perceiving empathy of reactions in a text transcription. Our model will be trained and evaluated with the proposed extension of the OMG-Empathy Prediction dataset. The original dataset is composed of 80 videos where an actor tells a story to a subject. Those stories were specifically designed to provoke different affective reactions in the subject. The subjects provided continuous self-assessment annotation on how they felt using a scale from positive (1) to negative (-1) values. To obtain all the necessary language information from this dataset, we extended the OMG-Empathy dataset with a manual transcription of all the videos. Mostly, the subject reacts to the story with non-verbal actions such as filler words (including uh, uh-huh, yeah, and hmm) and laughs, breath, and agreement. Those non-verbal reactions were cautiously annotated in the transcriptions. So, our proposed model comprises two steps: first, an RNN model is fed the actor's sentence and generates an overall feature map; then, we use this feature map with the sequence length, and the previous conversation feature map with an empathy predictor, trained with the OMG-Emotion train data. We evaluate our model using the Concordance Correlation Coefficient (CCC) between the model's prediction and the self-assessments of the users, using the same train-test division proposed in the OMG-Empathy dataset. We envision that our model will integrate short-term and contextual information to predict empathetic responses, allowing improved performance on human-robot interactions.