Ommie in the Home: A Non-Invasive Learning Model for Social Robotics

Grace Abawe; Advisor Brian Scassellati; Advisor Rajit Manohar

Introduction

Privacy in Social Robotics and Human Robot Interaction is essential in home-deployable robots. Importantly, sensors equipped by a robot determines the sensitivity of data that can be collected.

Ommie is a deep-breathing, anxiety reduction robot, meaning it's sociability should be prioritized. One robot behavior that increases personability is recognizing the intent to interact. However, for Ommie to be an in-home robot, it should preserve user privacy by limiting data collection that is user-identifiable. Rather than using invasive data collected by sensors like cameras, Ommie should use low-level sensors to achieve the same behaviors.

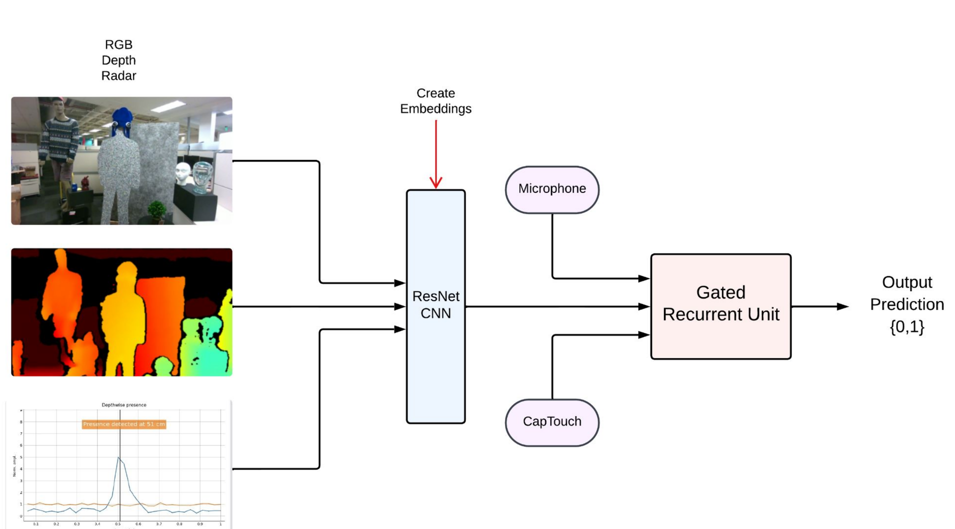
Thus, can Ommie leverage low-level sensors like a radar to predict a user's intent to interact while preserving user privacy?

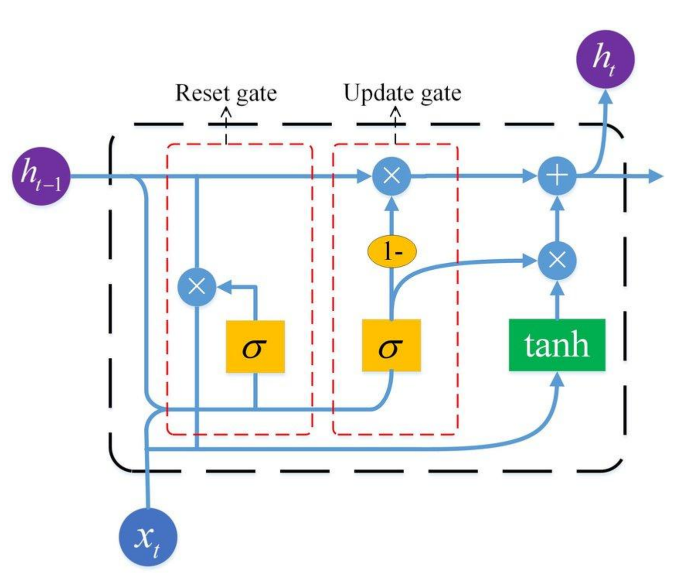
Machine Learning Model

Goal: Training a model to predict a user's intent to interact with the robot

Gated Recurrent Unit Recurrent Neural Network (GRU RNN)

- Good for multimodal datasets
  - High dimension data compressed by CNN into lower-dimension embeddings
- Processes sequential data (radar, audio)
- Uses gating mechanisms to selectively update the hidden state of the network at each time step
  - Reset gate: how much of the previous hidden state should be forgotten
  - Update gate: how much of the new input should be used to update the hidden state
- Captures long-term dependencies, detects subtle changes in behavior patterns over time





Methods

Sensors

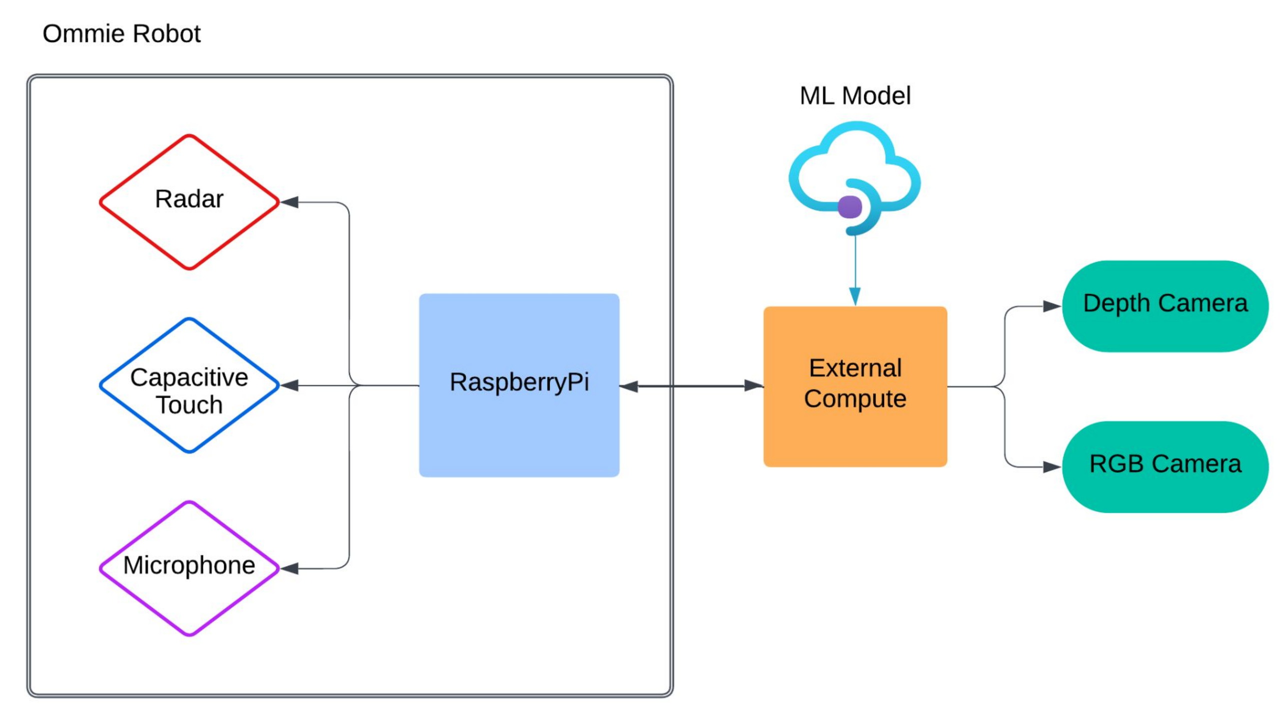
Wave Radar Sensor - Emits waves that are reflected by objects to calculate distance. Used for presence detection.

Multi-Directional Microphone - 4 microphone arrays that capture audio, decibel level, direction of arrival, and detect voice.

Capacitive Touch Sensor - Detects capacitive loads on touch contacts (i.e. the sensor is touched).

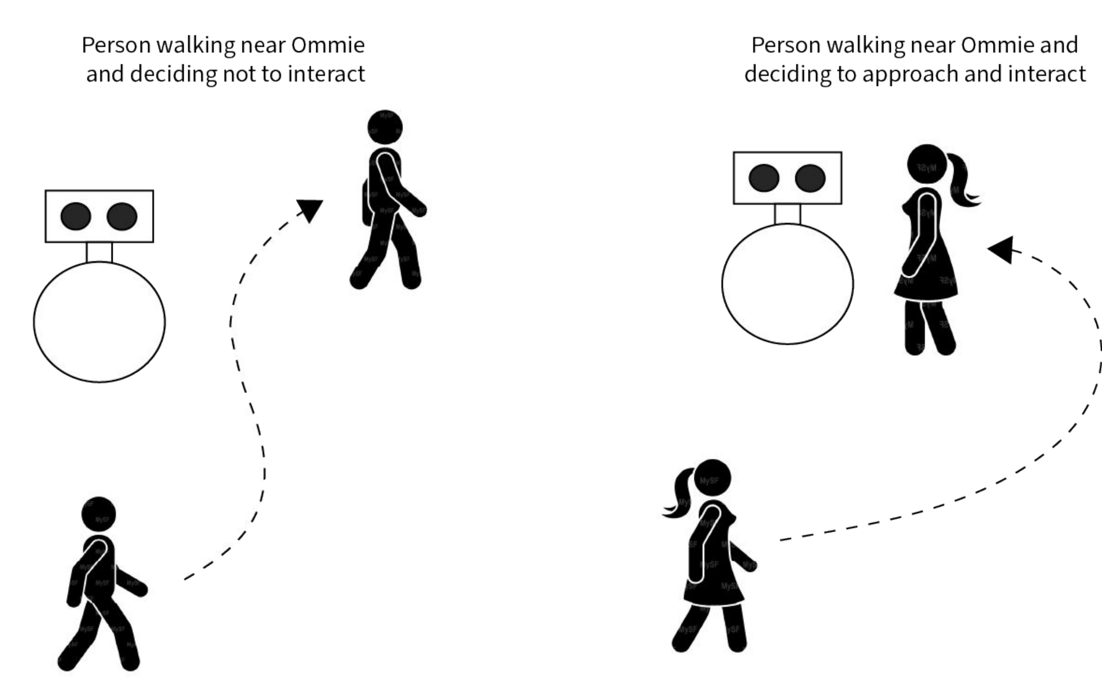
Camera - Records RGB and Depth images. Used as baseline for comparison.

Hardware Setup



Experiments

- Trials of people walking by Ommie and either:
  - Approaching and Interacting
  - Ignoring
- Interaction = Touching CapTouch
- Intent to Interact label recorded as 4 seconds before touch timestamp, else no interaction



Results

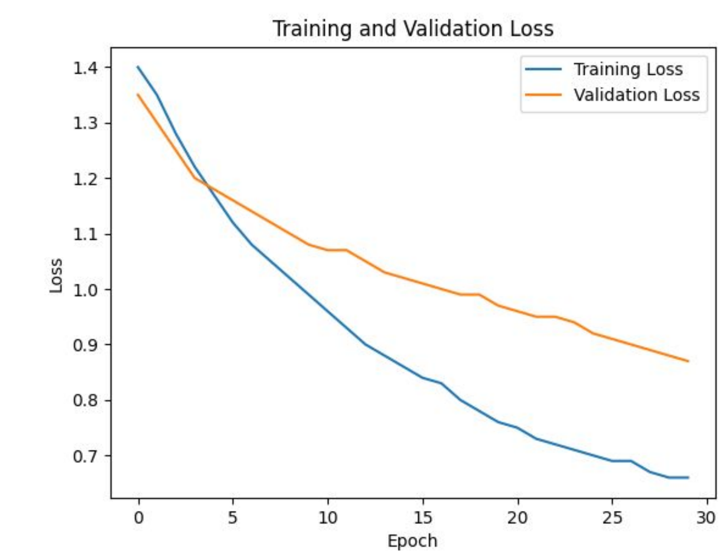


TABLE I
RESULTS OF MODEL PREDICTIONS.

Sensor	Accuracy	Precision	Recall	F1 Score
Radar + Microphone	0.43	0.33	0.37	0.35

Conclusions

The model correctly predicted the intent to interact 43% of the time. Based on the results of the model predictions, the use of only radar and microphone alone does produce inferior results compared to high-level sensors. However, it is also observable that it did not perform terribly in the context of predicting intent to interact, a nontrivial problem. The system also produced false positives more often than false negatives, meaning Ommie would greet users more often than necessary. However, for a user trying to preserve privacy, such a system may be acceptable in comparison to the use of methods like gaze detection. Consequently, further research investigating privacy and specific sensor usage, especially in the context of robots deployed in the home must be conducted.

Future Work

- Try other sensors such as infrared motion detector or lidar
- Run more trials to create larger dataset & train better model
- Produce other smart/sociable Ommie behaviors such as reminding users to do their breathing exercises
- Integrate system into actual Ommie robot and deploy in real life

Acknowledgements

A special thank you to my graduate mentor Fern Limprayoon.
Thank you to the Timothy Dwight Mellon Senior Forum Research Grant for their generous funding.