Augmented Reality and Affective Computing on the Edge Makes Social Robots Better Companions for Older Adults

Taif Anjum, Steven Lawrence and Amir Shabani
School of Computing, University of the Fraser Valley, British Columbia, Canada

Keywords: Affective Computing, Social Companion Robots, Socially-Assistive Robots, Augmented Reality, Edge Computing, Embedded Systems, Deep Learning, Age-Friendly Intendent Living, Smart Home Automation.

Abstract: The global aging population is increasing rapidly along with the demand for care that is restricted by the decreasing workforce. World Health Organization (WHO) suggests the development of smart, physical, social, and age-friendly environments will improve the quality of life for older adults. Social Companion Robots (SCRs) integrated with different sensing technologies such as vision, voice, and haptic that can communicate with other smart devices in the environment can allow for the development of advanced AI solutions towards an age-friendly, assistive smart space. Such robots require the ability to recognize and respond to human affect. This can be achieved through applications of affective computing such as emotion recognition through speech and vision. Performing such smart sensing using state-of-the-art technologies (i.e., Deep Learning) at the edge can be challenging for mobile robots due to limited computational power. We propose to address this challenge by off-loading the Deep Learning inference to edge hardware accelerators which can minimize the network latency and privacy/cybersecurity concerns of alternative cloud-based options. Additionally, to deploy SCRs in care-home facilities we require a platform for remote supervision, assistance, communication, and technical support. We propose the use of Augmented Reality (AR) smart glasses to establish such a central platform that will allow one single caregiver to assist multiple older adults remotely.

SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

The global aging population is increasing at a rate faster than ever. According to the World Health Organization (WHO), the world's aging population aged 60 years and older is expected to total 2 billion by 2050, up from 900 million in 2015 (Steverson, 2018). The demand for care is increasing while the supply is restricted due to the decreasing workforce. WHO suggests that the development of smart, physical, social, and age-friendly environments will improve the quality of life for older adults (Ionut, et al., 2020). To make the living space more personalized, connected and socially amenable, such environment could utilize advances in Artificial Intelligence (AI) and robotics such as Social Companion Robots (SCRs) (Mitchinson & Prescott, 2016)) and computer vision for scene understanding through human motion tracking (Ghaeminia, Shabani, & Shokouhi 2010), objects relationship (Shabani & Matsakis 2012) and monitoring human daily activities (Shabani, Clausi, & Zelek 2011- 2013) when integrated with smart home automations such as intelligent occupancy (Luppe & Shabani, 2017), and smart ventilation (Forest & Shabani, 2017).

Augmented Reality (AR) through smart glasses has become a widely popular multidisciplinary research field over recent years in a wide range of fields such as healthcare (Sheng, Saleha, & Younhyun, 2020), military, manufacturing, entertainment, games, educations, teleoperation and robotics (Varol, 2020). However, the literature lacks research of AR with SCRs for elderly care. AR gives the real-time view of our physical world with the addition of interactable computer generated objects. AR can be experienced with mobile displays, computer monitors and Head-Mounted Displays (HMDs). In the recent years, AR smart glasses such as Microsoft HoloLens and Google Glass allowed for efficient and realistic interaction between humans and autonomous systems. Among them Microsoft HoloLens 2 is one of the state-of-the-art commercially available device used in many applications (Xue, et al., 2020).

For health-care specific to older adults, AR related works focus on the Physical, Social and

Anjum, T., Lawrence, S. and Shabani, A.
Augmented Reality and Affective Computing on the Edge Makes Social Robots Better Companions for Older Adults.
DOI: 10.5220/0010717500003061
In *Proceedings of the 2nd International Conference on Robotics, Computer Vision and Intelligent Systems (ROBOVIS 2021)*, pages 196-204
ISBN: 978-989-758-537-1
Copyright © 2021 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

Psychological well-being. Physical well-being include training for their lack of ability, encouraging physical activity, providing reminders for health related activities (e.g., medicine or food intake) (Lee, Kim, & Hwang, 2019). For social well-being, AR offers remote participation, virtual interaction and emotional relationship for older adults who face decline in mobility, lacks transportation or has financial constraints (Lee, Kim, & Hwang, 2019). For physiological well-being, interactive games improve their moods and augmented immersive worlds offer an escape to forget their chronic pain, anxiety and social isolation (Lee, Kim, & Hwang, 2019).

AR has been used with a social robot for medical dose control (Lera F.J., 2014). It also has been used to engage the patients with dementia in a relaxing nature experience (Feng, Barakova, Yu, Hu, & Rauterberg, 2020). Other works include AR-based coaching, exercise games, e-learning for older adults (Lee, Kim, & Hwang, 2019). However, among the extensive research in this area, little work includes AR with social robots for older adult care.

Our application of SCRs is targeted towards care home facilities. One of our goals is to reduce the pressure on overworked staff and caregivers in care home facilities, allowing them to better focus on their most important person-to-person duties. We aim to design a system where the caregivers can instantly communicate to the older adult, provide assistance through actuation, analyze their clinical data, and perform remote intervention when necessary. To achieve this, we propose the use of Augmented Reality smart glasses (e.g., Microsoft HoloLens 2).

Integrating SCRs with AR smart glasses will allow for an easy-to-use central platform for monitoring, actuation, communication, assistance, and system troubleshooting. Typical monitoring systems in hospitals and care-home facilities rely on feeds from CCTV cameras that are monitored by technicians. When technicians notice events such as a fall, the caregivers are alerted. With this approach, the delay in receiving the alert can result in serious injuries. Other approaches rely on sensors to detect movements or changes in vitals and alert caregivers through mobile applications. After receiving the alert, the caregiver must locate the older adult to provide the support.

For our application of deploying a group of SCRs in care-home facilities, the number of caregiver could be as low as one person who monitors all SCRs and older adults through a central platform over AR interface. With our approach, the duties of the middleman (technician) will be replaced by the automation and the caregiver will have direct access for intervention. When comparing a web-based or mobile application to our AR-based system, accessing the platform through smart glasses allows for a more convenient, portable, and immersive experience where the caregiver can remain engaged in their daily duties. Furthermore, using a mobile device instead of smart glasses can be challenging as using mobile devices in workplaces are controversial. With the increasing use of AR technologies, soon having smart-glasses instead of mobile phones may become the norm. Particularly in healthcare, the use of AR is becoming increasingly popular. The next-generation smart glasses are estimated to be reduced in size to be comparable to standard eyeglasses that will allow such integration with ease.

Although a platform for supervision, health monitoring and communication is essential for older adult care, that is just one component of our vision of SCRs integrated with smart devices. What differentiates SCR with other assistive technologies is the interaction component. The robot needs to interact with the human counterpart in a natural human-like manner. More specifically, it needs to recognize and respond to human emotion. Applications of Affective Computing can allow social SCRs to achieve this ability. Affective Computing allows systems and devices to recognize and respond to human affect (e.g., emotion, touch). The main contributions of this paper are as follows:

1) Integration of SCRs (i.e., Miro-e) with AR Smart Glasses (i.e., HoloLens 2) for supervision and support in care home facilities. This integration enables interoperation while utilizing the benefits of both systems.

2) Hardware integration of Miro-e's Raspberry Pi 3 (B+) to Nvidia's Jetson Nano for off-loading Deep Learning-based Facial Emotion Recognition on the edge. Our choice of integration using wired ethernet minimizes the network latency and privacy/cybersecurity concerns.

The rest of the paper is organized as follows. Section 2 explains our proposed framework of using SCR and AR and their integration, Section 3 presents a computation off-loading method for performing Deep Learning inferences on the edge. To test and demonstrate our method we deploy FER model into our robot Miro-e. Section 4 explains the integration of HoloLens 2 with Miro-e for a central supervision, communication and actuation platform for older adult care. Section 5 presents the conclusion and future work.

2 PROPOSED FRAMEWORK

Our proposed system is to deploy SCRs in different elder's room for individual monitoring and personalized interactions through speech and vision. One could consider SCR robot as a limited version of a private all-time available nurse. To address the privacy concerns, as a standalone system, each SCR could analyse the sensitive data such video/speech recording and only communicate processed and aggregated data to the central interface on an AR which is accessed by the caregiver. To address the cloud/internet cybersecurity concerns, a private local area network could be setup in the facility to connect the SCRs to the central interface. As it can be seen in Figure 1, the caregiver with AR system can monitor the analysed status of the elders and provide necessary support through the SCR without needing to be present in the elder's room. This enables the caregiver who wears the AR system to effectively and efficiently interact with multiple people through their dedicated SCR robot from one (remote) location. The central interface shows different smart rooms of the older adults with SCRs integrated with other smart devices in the rooms. Through different interactive screens, the interface can provide different analysed information such as overall emotional states of the older adults in each room over a period of time. In case the caregiver notices unusual behaviour they can take appropriate actions. For example, the caregiver notices that the older adult in room 3 was sad 80% of the time within a period. They can initiate a video/voice call with the older adult to hear their concerns, help improve their mood, check their vitals to ensure they are healthy and even dispatch for extra support and alert friends and family.

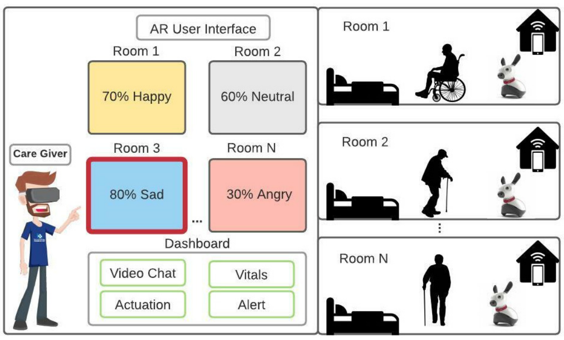

Figure 1: Our proposed AR-based system integrated with Social Companion Robot for older adult care.

For our application, emotion recognition through speech and vision can make the experience more interactive and engaging while providing feedback on the user's mental health and well-being. With the exponential increase of data, deep learning techniques are being more widely used due to their superiority in performance compared with the conventional machine learning techniques, especially when large amounts of data is available. However, performing deep learning training and inferences in embedded systems such as SCR robot is challenging due to the computational cost of deep learning algorithms. In such cases, researchers utilize the cloud to perform the computation. But cloud based approaches incur large latency, energy and financial overheads and also privacy/cybersecurity concerns (Mittal, 2019). In-order to mitigate these challenges, edge computing is used in the literature, that means computation is performed where the data is produced (or close to it). To meet the computational needs of deep learning algorithms on the edge, multiple companies launched low-power hardware accelerators (Mittal, 2019). Among them, NVDIA's Jetson is one of the most widely used for Deep Learning inference. Jetson features CPU-GPU heterogenous architecture where the CPU can boot the OS and the CUDA-capable GPU perform complex machine-learning tasks (Mittal, 2019). To minimize the network latency and privacy/cybersecurity concerns, in-order to perform Deep Learning inferences on the edge, we propose to off-load the computation to Nvidia's Jetson through a hardware integration via ethernet cable. To test our system we deployed our Facial Emotion Recognition (FER) model where the computation was performed by Jetson using data from Miro-e's cameras.

3 REAL-TIME FER ON THE EDGE

Facial Emotion Recognition (FER) has been a topic of interest in the computer vision community for many applications. Specifically for Affective computing, FER is an integral part of affect recognition. Standard ML algorithms such as SVMs and their variations have been extensively used for FER classification. Over the recent years, Deep Learning techniques with Convolutional Neural Networks (CNNs) have proven to outperform standard ML algorithms for image and video classification in FER (Lawrence, Anjum, & Shabani, 2021). However, Deep Learning training and inferences are computationally expensive and typically performed using powerful and expensive computers or servers. Most mobile robots such as Miro-e have limited on-board resources such as processor, memory and battery. To apply Deep Learning algorithms on such robots', researchers typically off-load heavy computations to cloud hosts. The data is collected from the mobile robot and sent to the cloud for applying Deep Learning algorithms. But this makes the data vulnerable to cybersecurity concerns and privacy invasion due to data transmission through computer networks (Chunlei, et al., 2020). Additionally, network failure or network package loss can disrupt cloud-based deep learning. (Chunlei, et al., 2020). In order to apply Deep Learning algorithms on mobile robots we require hardware that are energy efficient, small in size and affordable (Mittal, 2019). Furthermore, for our application we require edge computing for its low latency and data privacy since sensitive data is processed on-board and not on the cloud (Mittal, 2019).

To satisfy the need for Deep Learning inference on the edge, serval products have been launched by commercial vendors based on hardware accelerators. Apart from Graphics Processing Units (GPUs), system-on-chip architectures that utilize the power of Application-Specific Integrated Circuits (ASICs), Field-programmable Gate Arrays (FPGAs) and Vision Processing Units (VPUs) can also be used for inference at the edge (Amanatiadis & Faniadis, 2020). Some of the most widely used commercially available devices include (Amanatiadis & Faniadis, 2020):

- The Edge TPU by Google, is an ASIC exclusively for inference achieving 4 Tera Operations Per Second (TOPS) for 8-bit integer inference. However, it requires the models to be trained using TensorFlow which is a limitation. Google's Coral Dev Board featuring Edge TPU is priced at $169 (USD) for the 4GB RAM version.

- The Intel Neural Compute Stick 2 is a System-on-Chip built on the Myriad X VPU, optimized for computer vision with dedicated neural compute engine for hardware acceleration of deep neural network inferences. It has max performance of 4 TOPS, similar to the Edge TPU. However, it requires a host PC since the device is distributed as a USB 3.0 stick. It also requires the model to be converted to a Intermediate Representation (IR) format that can slow down the development process. This device is priced at $68.95 (USD) excluding the cost of a host PC.

- NVIDIA's Jetson series is a group of embedded machine learning platforms that aims to be computationally powerful while being energy efficient. They feature CUDA-capable GPUs for efficient machine learning inferences. Their cheapest and lightest model is the Jetson Nano TX1 with a peak performance of 512 single precision (SP) Giga floating-point operations per second (GFLOPS). They cost $129 (USD) for the 4GB version.

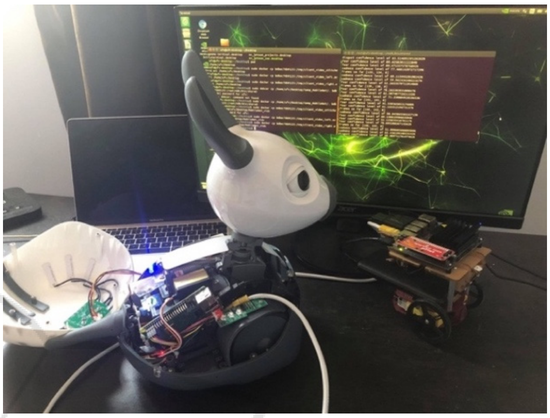

Figure 2: Hardware integration of Miro-e and Nvidia's Jetson Nano via ethernet port.

Unlike the EDGE TPU and Neural Compute Stick 2, Jetson Nano does not require a host PC, models do not require any conversion, is not restricted to any specific machine learning framework and is the cheapest option. However, Jetson Nano has a lower computational power compared to the other two devices, but the difference will not be significant for our applications of vision and speech inferences. Considering all the components, Jetson Nano was the best fit for our system. Figure 2 shows our hardware integration of Miro-e and Nvidia's Jetson Nano for efficient Deep Learning inference at the edge. In the image, Jetson Nano board is attached to a robot body. For our final prototype we will have the Jetson Nano board on Miro-e's body.

Table 1 shows the difference in power between Jetson Nano TX1 and Raspberry Pi 3(B+) (Miro-e's on-board computer). The Jetson Nano TX1 significantly outperforms Miro-e's on-board computer in every aspect. Having such a powerful computer integrated with Miro-e opens doors for a large number of applications.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

Table 1: Jetson Nano and Miro-e's on-board computer specification comparison.

	Jetson Nano TX1	Raspberry Pi 3(B+)
GPU	256-core Maxwell @ 998MHz	VideoCore IV
CPU	ARM Cortex-A57 Quad-Core @ 1.73GHz	ARM Cortex-A53 Quad-Core @ 1.4GHz
Memory	4GB 64-bit LPDDR4 @ 1600MHz, 25.6 GB/s	1GB LPDDR2 @ 900MHz, 8.5 GB/s
Peak Performance	512 SP Gflops	6 DP Gflops

3.1 CNN Architecture for FER

In our recent studies, we introduced a data augmentation technique for FER using face aging augmentation. Publicly available FER datasets were age-biased. To increase the age diversity of existing FER datasets we used GAN based face aging augmentation technique to include representation of our target age group (older adults). We conducted comprehensive experiments for both intra-dataset (Lawrence, Anjum, & Shabani 2021) and cross-dataset (Anjum, Lawrence, & Shabani 2021) that suggest face aging augmentation significantly improves FER accuracy.

For the FER implementation, we utilize two Deep Learning architectures, CNNs in particular have shown great promise for image classification. For the purpose of FER, several studies showed that CNNs outperform other state-of-the-art methods. For our experiments, we used two CNN architectures; MobileNet, a lightweight CNN developed by Google and a simple CNN which we will refer to as Deep CNN (DCNN). (Howard, et al., 2017). Our DCNN classifier includes six convolutional 2D layers, three max-pooling 2D layers, and two fully connected (FC) layers. The Exponential Linear Unit (ELU) activation function is used for all the layers. The output layer (FC) has nodes equal to the number of classes (in this case, six classes) with a Softmax activation function. To avoid overfitting, Batch Normalization (BN) was used after every convolutional 2D layer and dropouts were used after every max pooling layer. Both BN and dropout were used after the first FC layer.

Additionally, we used a lightweight CNN architecture known as MobileNet. A lightweight model is required for our application of FER in embedded systems at the edge such as SCRS. MobileNet has 14 convolutional layers, 13 depth wise convolutional layers, one average pooling layer, a FC layer and a output layer with the Softmax activation function. BN and Rectified Linear Unit (ReLU) are applied after each convolution. MobileNet is faster than many popular CNN architectures such as AlexNet, GoogleNet, VGG16, and SqueezeNet while having similar or higher accuracy. The main difference between DCNN and MobileNet is that the latter classifier leverages transfer learning by using pre-trained weights from ImageNet. Throughout every experiment MobileNet contained 15 frozen layers from ImageNet. An output layer was added with nodes equal to the number of classes and softmax is used as the activation function. We used both DCNN and MobileNet classifiers with implementations from (Sharma, 2020). The Nadam optimizer was used along with two callbacks, 'early stopping' to avoid overfitting. For reducing the learning rate when learning stops improving, we used 'ReduceLRonPlateau'. The data is normalized prior to being fed into the neural networks as neural networks are sensitive to un-normalized data. Both the models we deployed onto Miro-e (MobileNet and DCNN) were trained with both original and face age augmented images.

3.2 Miro-e and Jetson Nano Integration: Off-Loading Deep Learning Inference to an Edge Computing Device

In order to established a connection between Miro-e and Jetson Nano we utilized the Robot Operating System (ROS). The connection could be established either through Wi-Fi or hardwired via an ethernet cable. As privacy is our top most priority we decided to go with hardware integration to keep Miro-e disconnected from the internet.

For the integration to be successful to we had to setup the environment such that Jetson and Miro-e were compatible. Miro-e requires ROS Noetic which is the latest version of ROS that is only compatible with Ubuntu 20.04 Focal Fossa. However, Jetson does not support Ubuntu 20.04 so we had to use a Docker image on Ubuntu 18.04 with ROS noetic built from source. Then we had to install TensorFlow, OpenCV and other Python and ROS libraries. We also installed Miro Development Kit (MDK) on to the Jetson Nano and established a connection through ethernet port. The communication was done through ROS libraries. We were successfully able to perform real-time FER with both MobileNet and DCNN models without any lag or disturbances.

Miro-e is equipped with various sensors including one camera in each eye. Each sensor is recognized as a topic in the ROS interface. In ROS, topics are named buses over which nodes can exchange messages. The left eye camera and right eye camera topics allowed us to view Miro's real-time video feed and utilize each frame to detect and track faces. Once the face is detected, is it then cropped, re-sized to 48x48, converted to grayscale and passed onto our trained FER model for emotion prediction. This entire process is done using one script. Figure 3 demonstrates our real-time FER through Miro-e.

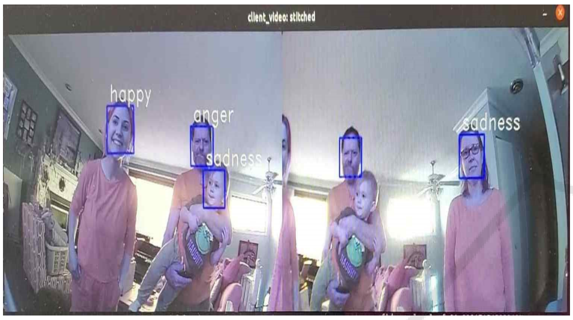

Figure 3: Real-time FER with Miro-e on the edge.

4 AR SMART GLASS INTEGRATION WITH SCR

Our proposed AR-based system for monitoring group of social robots in care home facilities requires deploying a group of SCRs with the following constraints:

1) Controlling robots to assist the elderly while ensuring they performing as expected.

2) Running diagnostics on such robots in case of hardware or software issues can be expensive.

3) Non-intrusive monitoring and communication with caregivers.

Having a central monitoring platform for a group SCRs is imperative for our application. The platform will be used to monitor the health and well-being related data from various sensors in the environment as well other sensing data from vision, speech, haptic (e.g., emotion recognition). A video/voice chat option enables the caregiver to interact with the older adult regarding concerns about their health and well-being (e.g., Miro-e detects the older adult is sad or angry). The system recommends intervention when Miro-e detects using various sensors in the environment that the older adult has skipped medication, food or exercise. The platform will further reduce the risk on contamination during viral outbreaks such as COVID-19.

4.1 Connection between ROS and Unity

Figure 4 outlines the proposed system architecture where the left side represents the user interface of Microsoft HoloLens 2 with Unity game engine. This requires the ROS Noetic running on a Ubuntu 18.04 (Bionic Beaver).

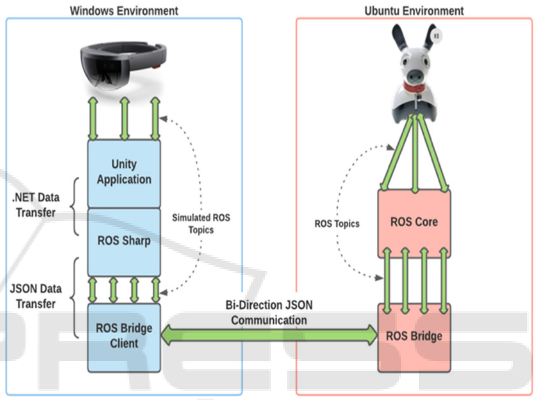

Figure 4: HoloLens-Miro end-to-end System Architecture.

Table 2: System Development tools.

Development Stack
Microsoft Visual Studio 2019
Unity
ROS-Sharp
Mixed Reality Toolkit
ROS-bridge suite
Miro Development Kit
Client Stack
HoloLens 2
Miro-e
Server Stack
Ubuntu 18.04 Bionic Beaver
ROS Noetic

The development tools are presented in Table 2. In order to establish the communication between Miro-e and Microsoft HoloLens 2 we utilized Unity, ROS Sharp and RosBridge-suite. ROS Sharp is a set of open source software libraries and tools in C# for communicating with ROS from .NET applications, particularly Unity. RosBridge-suite is a collection of packages including RosBridge and RosBridge

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

WebSocket. RosBridge is a .NET JSON API for communication between ROS and non-ROS programs. RosBridge WebSocket allows messages to be exchanged between ROS nodes.

The Ubuntu machine runs Roscore and RosBridge. Roscore is a collection of nodes and programs that are pre-requisites of a ROS-based system. Roscore must be running in order for ROS nodes to communicate. Roscore is used to send data through topics to Miro-e, RosBridge is used to accept data from Unity through the RosBridge WebSocket. The data is then published on ROS topics for Miro-e to access the data.

Unity sends simulated ROS message types to the WebSocket provided by the RosBridge client hosted on the Ubuntu machine. The RosBridge client then translates these simulated messages from unity to ROS and publishes them to ROS topics.

The first step is to establish a connection between ROS and Unity project. To accomplish this, we created a Unity project and copied RosSharp and NewtonSoft (JSON framework for .NET) into the project. We then installed and configured the Mixed Reality Toolkit (MRTK), and configured RosBridgeClient to be used with Unity project. This setup establishes a connection between ROS master node and the Unity project, allowing bi-directional messages to be sent or received between ROS and the Unity project.

4.2 Building an AR HoloLens App

The AR application is created using the Unity game engine. To exchange messages with HoloLens 2, we require an Universal Windows Platform (UWP) app to run in HoloLens 2. For this, we first developed the app in Unity using C# and then deployed it to HoloLens 2. We used the Mixed Reality Tool Kit's (MRTK) button and menu prefabs alongside Unity's TextMeshPro to build a simple user interface (UI). The user interface has buttons with the corresponding functions written on it such as red LED, blue LED, wag tail and so on. The predicted emotion from our FER model is displayed on the top of the UI.

4.3 Publishing to ROS Topics

ROS topics are named buses over which nodes can exchange messages. Each ROS topic is constrained by the ROS message type used to publish to it and nodes can only receive messages with a matching type. Both Miro and the Unity project can have their own topics and we can also build custom topics. The Unity project is required to be subscribed to the specific Miro topics from/to which it intends to exchange messages and vice-versa. Miro-e has various topics including illumination, kinematic joints, cosmetic joints, mics, camera left, camera right, etc. We can publish a message (executable code) to these topics triggering the corresponding functions to execute allowing Miro to move, wag tail, light-up etc. To begin this process, we first initialize Roscore with the master node as the Ubuntu machine. Then we connect the RosBridgeClient Node and Miro's ROS node to the master node. Once both nodes are linked to the master node we are ready to exchange messages between Miro and our app in Unity. We use a script that allows Miro to listen for data to be published to the subscribed topics. Once Miro's ROS node receives a publish request it executes the command for that specific topic. For example, we send a request to execute a block of code to Miro's Illumination topic from ROS Sharp in Unity. As ROS Sharp is subscribed to Miro's Illumination topic, it can exchange messages with that topic. Miro is constantly listening for publish requests and accepts a request that matches the type of the topic. A similar process takes place when executing commands to HoloLens 2, a publish request is sent to ROS Sharp topics in Unity from Miro's ROS node.

Figure 5 is a picture of our prototype. Miro-e performs FER and sends the predicted emotion to the HoloLens AR app. The emotion is then displayed on the user interface. Furthermore, using the user interface, the caregiver can manipulate Miro-e to navigate, change LED light colors, wag its tail, move its head, ears and perform every function it is capable of.

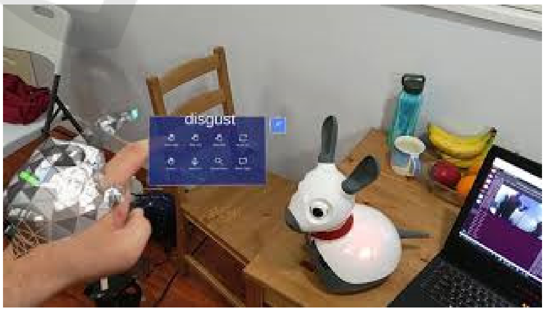

Figure 5: Prototype running on HoloLens 2.

5 CONCLUSION

Given the rise of global aging population it is imperative to develop systems for age-friendly smart environments to support our aging population for both independent living and also in long-term care facilities. Integrating SCRs to such smart environments can improve the quality of life for older adults. With the use of personalized machine learning, SCRs can learn the preferences of the older adults such as preferred temperature, lighting intensity, and even activate robotic vacuum cleaners (e.g., Roomba (Tribelhorn & Dodds, 2007) according the their preferred time.

We proposed an AR-based system for interactive interfacing with multiple SCRs deployed in different rooms/homes. More specifically, we developed a seamless communication between Microsoft HoloLens 2 and Miro-e robot. This integration serves as an efficient platform for controlling the robot for assisting the older adult, over-ride control in case the robot takes unexpected actions, monitor their health and daily activities (i.e., medication or food intake, exercise), instant communication for emergency situations and much more.

For a more natural interactions between the robot and elder, we developed an improved deep learning based facial emotion recognition technique for affective computing. To overcome the limited computational power of the robot's computer, we off-loaded the Deep Learning inference to on the edge hardware accelerators which opens doors for a wide range of applications. To achieve this we integrated Miro-e to Nvidia's Jetson and successfully performed our FER algorithms on the edge and minimize the network latency and privacy/cybersecurity concerns of alternative options which require cloud and internet connectivity.

Having a central interactive platform through AR smart glasses for managing multiple robots and being able apply state-of-the-art learning algorithms on the edge is a milestone towards deployment of SCRs in smart environments to assist older adults. In fact, combining our proposed AR-based system with applications of Affective Computing allows for a more reliable and safer interaction between the SCR and the older adult.

Our future work will include integration of Miro-e with smart home devices for advanced personalized home automation for elder adult. The integration will focus on the safety, security, lighting and heating/cooling control, and also mental health of the older adult. Another research direction is study of human factors in the interface design and adding more functionalities for communication, alert and health analysis to the interface. Another next step is the field study to assess the usability, acceptance rate, and benefits of our systems.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contribution of the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Interactive Intelligent Systems and Computing research group through which this project is supported.

REFERENCES

Amanatiadis, E., & Faniadis, A. (2020). Deep Learning Inference at the Edge for Mobile and Aerial Robotics. *IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)* (pp. 334-340). IEEE.

Anjum, T., Lawrence, S.& Shabani, A. (2021). Efficient Data Augmentation within Deep Learning Framework to Improve Cross-Dataset Facial Emotion Recognition. *The 25th Int'l Conf on Image Processing, Computer Vision, & Pattern Recognition.* Las Vegas: Springer Nature.

Chunlei, C., Peng, Z., Huixiang, Z., Jiangyan, D., Yugen, Y., Huihui, Z., & Zhang, Y. (2020). Deep Learning on Computational-Resource-Limited Platforms: A Survey. *Mobile Information Systems* (pp. 1-19). Hindawi .

Forest, A., Shabani, A. (2017). A Novel Approach in Smart Ventilation Using Wirelessly Controllable Dampers. Canadian Conference on Electrical and Computer Eng. Windsor, Canada: IEEE.

Feng, Y., Barakova, E., Yu, S., Hu, J., & Rauterberg, G. (2020). Effects of the Level of Interactivity of a Social Robot and the Response of the Augmented Reality Display in Contextual Interactions of People with Dementia. *Sensors*, (pp. 20(13), 3771).

Ghaeminia, M.H., Shabani, A.H., & Shokouhi, S.B. (2010)/ Adaptive motion model for human tracking using Particle Filter. International Conference on Pattern Recognition, Istanbul, Turkey.

Howard, A., Menglong, Z., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861.*

Ionut, A., Tudor, C., Moldovan, D., Antal, M., Pop, C. D., Salomie, I., Chifu, V. R. (2020). Smart Environments and Social Robots for Age-Friendly Integrated Care Services. *International Journal of Environmental Research and Public Health.*

Keven, T. K., Domenico, P., Federica, S., & Philip, W. (2018). Key challenges for developing a Socially Assistive Robotic (SAR) solution for the health sector. *23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD).* Romee, Italy: IEEE.

Lawrence, S., Anjum, T., & Shabani, A. (2021). Improved Deep Convolutional Neural Network with Age Augmentation for Facial Emotion Recognition in

Social Companion Robotics. *Journal of Computational Vision and Imaging Systems* (pp. 1-5). JCVIS.

Lee, L., Kim, M., & Hwang, W. (2019). Potential of Augmented Reality and Virtual Reality Technologies to Promote Wellbeing in Older Adults. *Applied Sciences*, 9(17), 3556.

Lera F.J., R. V. (2014). Augmented Reality in Robotic Assistance for the Elderly. *González Alonso I. (eds) International Technology Robotics Applications. Intelligent Systems, Control and Automation: Science and Engineering.* Springer, Cham.

Luppe, C., & Shabani, A. (2017). Towards Reliable Intelligent Occupancy Detection for Smart Building Applications. Canadian Conference on Electrical and Computer Eng., Windsor, Canada: IEEE

Mitchinson, B., & Prescott, T. J. (2016). MIRO: A Robot 'Mammal' with a Biomimetic Brain- Based Control System., (pp. 179-191).

Mittal, S. (2019). A Survey on optimized implementation of deep learning models on the NVIDIA Jetson platform.

Rokhsaritalemi, S. A.-N.-M. (2020). A Review on Mixed Reality: Current Trends, Challenges and Prospects. *MDPI*, 636.

Shabani, A.H., Clausi, D.A., & Zelek, J.S. (2011), Improved spatio-temporal salient feature detection for action recognition. British Machine Vision Conference. Dundee, UK: BMVC.

Shabani, A.H., Clausi, D.A., & Zelek, J.S. (2012), Evaluation of local spatio-temporal salient feature detectors for human action recognition. Canadian Conf. On Computer and Robot Vision, Toronto, Canada: IEEE,

Shabani, A.H., Zelek, J.S., & Clausi, D.A. (2013), Multiple scale-specific representations for improved action classification. Journal of Pattern Recognition Letters.

Shabani, A.H., Ghaeminia, M.H., & Shokouhi, S.B. (2010). Human tracking using spatialized multi-level histogramming and mean shift. Canadian Conf. on Computer and Robot Vision. Ottawa, Canada: IEEE.

Shabani, A.H., & Matsakis, P. (2012). Efficient computation of objects' spatial relations in digital images. Canadian Conf. On Electrical and Computer Eng. Montreal, Canada: IEEE.

Sharma, G. (2020, April). *Facial Emotion Recognition*. Retrieved from Kaggle: https://www.kaggle.com/gauravsharma99/facial-emotion-recognition

Sheng, B., Saleha, M., & Younhyun, J. (2020). Virtual and augmented reality in medicine. In *In Biomedical Engineering, Biomedical Information Technology (Second Edition)* (pp. 673-686).

Steverson, M. (2018). *Ageing and Health*. Retrieved from https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

Tribelhorn, B., & Dodds, Z. (2007). Evaluating the Roomba: A low-cost, ubiquitous platform for robotics research and education. *In Proceedings 2007 IEEE International Conference on Robotics and Automation* (pp. 1393-1399). IEEE.

Varol, Z. M. (2020). Augmented Reality for Robotics: A Review. *MDPI*.

Xue, C., Shamaine, E., Qiao, Y., Henry, J., McNevin, K., & Murray, N. (2020). RoSTAR: ROS-based Telerobotic Control via Augmented Reality. *RoSTAR: ROS-based Telerobotic Control via Augmented Reality* (pp. 1-6). IEEE.