RESEARCH AND ANALYSIS OF CURRENT ROBOTIC AFFECTIVE PETS AND THEIR APPLICATIONS FOR CHILD THERAPY AND EDUCATION WITH SUGGESTED DESIGN IMPROVEMENTS

MARCH 31, 2022

UNIVERSITY OF WARWICK
Téa Clark

Contents

Chapter 1

Introduction

Aim

The aim of the project is to assist in providing an affordable alternative to social robots currently available on the market, specifically targeted towards enhancing the learning environment of children (particularly those with learning difficulties), reducing stress and being a tool for child psychologists, monitoring physiological signals and acting as a means of communication. It should provide an emotional support means between that of soft toys and therapy animals, giving an interactive experience without the hygiene issues, animal welfare and other logistical problems that come with having a live animal, as noted by R.Kimura et al. [1] and Clara Moerman et al. [2].

Background Information

There is a great deal of research towards the benefits of social robots and its application towards many areas of society including the elderly and children. In particular interest to the aim of this project, children's social interaction abilities have been reported as greatly improved by robot assistance in teaching/therapy sessions [3], students' interest in learning is greater [4], and their wellbeing is positively affected [2], with reduced stress levels following robot interactions [5]. It is suggested that robots act as a middle ground between inanimate toys and animate humans, potentially causing children with autism spectrum disorders (ASD) to elicit greater social behaviour [6].

The literature does show a lack of development of soft-bodied social robots within the education sector compared to other areas. Robots like Nao [7] and Pomelo [8] form the norm for social robots in a teaching environment, acting in a teaching assistant role rather than assisted therapy. In contrast, the standard for social robots for the elderly included much cuddlier robots like Paro [9] and Hasbro's Joy-for-all range [10]. Since the aim of this project includes relieving stress in a therapy setting, and forms of touching like stroking and hugging have been reported to reduce mental stress [11], this project will be taking inspiration from the health and social care industry-associated therapy robots and using a soft bodied form to encourage the use of touch from children, like traditional soft toys.

There is also a notable trend towards expensive therapy robots, with robots like Paro costing $6000 to buy outright [9] and Keepon $30,000 [12]. This is due to the quality of sensors and cameras used in the production, as well as the cost of medical approval. This project will be looking at creating an affordable alternative to these products, implying trade-offs will have to made during the hardware/software selection phase. It’s noted by Sartorato et al. that this lack of commercially available affordable options hinders the efficacy of social robot therapies, holding back an otherwise effective tool [13].

Objectives

The main objectives are to provide a summary of the existing literature surrounding social robots, specifically affective pets , application of this knowledge in the form of a design proposal and an analysis of the concept and suggestions for future improvement.

The research phase covers the benefits and reasons behind the social robots involvements in the social care sector (with a focus on children in particular); the existing features and capabilities of available robots; and the ideal behaviour and design of a robotic therapy pet. There will should also be an investigation into appropriate hardware and software for executing the desired functions of the “pet”.

From a background literature review the following functions have been suggested as features of interest:

1. Creating facial expressions and sounds
2. Remote control capabilities
3. Embrace recognition
4. Facial expression recognition
5. Voice recognition

They represent the most common features discussed by literature for achieving the aim outlined above and are present in current successful social robots available such as AIBO [14], Paro [9], Pleo [15] and Keepon [16].

The design concept aims to feature as many of the above features as possible, including the recognition system hardware, software, and a physical body design for the “pet”, within the time constraints of the project. A “pet” design capable of creating facial expressions and sounds represents the minimum requirement for the design, as this would allow a therapist/teacher to communicate indirectly with a patient, such that the child interacts with the teacher/therapist through the face of the robot, as with Keepon [16]. This would be enough to satisfy the aim of the target, given research suggests the presence of social robots increases the typical number of social interactions initiated by children with ASD [13], reduced stress levels [5] and generally improves the wellbeing of the user [2].

Additional features would increase versatility by allowing independent interaction to take place between “pet” and user, and also allow for improved conveyance of emotion from the user to a therapist, since robotic animals are suggested to be a “window into [a person’s] emotional state” [17].

The analysis of the concept will be covered in following sections of this report, namely Chapter 5 and beyond. It will cover a comparison between the project aims and design concept produced as well as suggestions for improvements and future work.

Chapter 2

Literature Review of Existing Social Robots

Benefits of social robots and affective pets

Animal assisted therapy (AAT) is already being implemented in a medical improvement, representing a non-drug intervention method with the ability to reduce pain in children, help manage children's behavioural stress and influence numerous other physiological factors, including oxygen saturation, heart rate, respiratory rate, diastolic blood press and cerebral oxygenation. [18] However, despite AAT's numerous benefits, it does carry some risks and logistical issues. For example, animals can carry diseases, harbouring harmful bacteria and pests on their bodies which is not conducive with keeping a sterile medical environment. It also can be harmful to an animal's welfare, potentially causing distress and putting the animal at risk of harm. [1] Robot assisted therapy provides an alternative – bringing the benefits of AAT whilst mitigating the risks.

It is noted by Kimura et al. that robot pets can provide a degree of social companionship and give some emotional satisfaction to a patient, offering an appropriate alternative to a companion animal. A robot pet could improve a child's mental wellbeing, providing mental stimulation, interaction and companionship independent of human intervention, helping to mitigate a shortage of nursing staff available at welfare facilities. [19] Moerman et al. also concur that social pet robots can provide similar benefits to play materials and animals, following an experiment demonstrating that interactions with a robotic pet can illicit positive emotions in children who have been subjected to stress [2].

The positive effects of robot assisted therapy for children are supported by numerous studies which note that robots can help encourage children to engage with therapy more positively [20] with approximately two-thirds of trials reporting positive results from incorporating social robots in therapy with people with ASD [21]. Jaishankar Bharatharaj, et al. notes that their study indicated "significant improvements in children's interaction abilities" when social robots were used to help deliver therapy sessions to children with ASD as opposed to sessions involving only humans [3]. The trend in interaction abilities is also continued by Kim et al. who state that in general children produce more speech and direct more speech towards adults (in the room) when in the presence of a social robot [22]. Kimura et al. suggests that robotic assisted activities are effective in stimulating conversion by making new topics, helping to reduce the build up of stress and boredom associated with time spent in hospital and medical treatments [1]. Robots have been shown to heighten engagement, decrease social anxiety and increase attention when used as therapeutic tools for children with austism spectrum disorders. It is suggested by Sartorato et al. that robotic interactions generate less frustration in individuals with ASD due to being more controllable, simplistic and predictable, thus being easier to interpret and respond to than human social interaction. Increased co-operation, sharing and turn taking behaviours, motivational and attentional engagement as well as reduction in repetitive behaviours and restricted interests are all listed as notable positive outcomes from robot-assisted therapy. [13]

Successes of particular robots include the utilisation of ‘KiliRo’ robot (a parrot-like design), who achieved notable learning improvements and reduction of stress levels (measured through urinary and salivary tests) in children with autism [5] and also Hugvies – a large huggable social robot which was shown to relieve emotional problems in listening situations and improve concentration and whose performance extended over a three-month period providing positive indications of long-term use effectiveness [23].

Physical Design

Many different physical forms of social robots can be seen in literature which can be broadly categorised into humanoid and non-humanoid. In a learning environment a human form was used most typically, however in a medical, therapy environment, animal forms seemed most prevalent. For example, a study by Kimura et al. into the use of robotic assisted activities for children in hospitals used a wide selection of social robots, including two cat-like robots and both dog and cat like interactive stuffed toys which importantly all fall under the animal umbrella [1]. It’s noted by Ricks and Colton that the most attentional engagement is elicited by non-humanoid robots [24] [13] and animal forms are concluded by Sartorato et al. to serve as effective social facilitators for application in communication and social therapies [13]. Due to the aim of the project directed towards designing a social robot for use in therapy, the above information has been used to narrow the scope of physical design down to a robotic pet/animal. Common forms are discussed below.

Paro [9] is a harp seal robot with soft padding and fur which is praised for feeling “robust” by stakeholders in the health and social are industry, reducing worries surrounding possible damage Results from the study conducted by Bradwell et al. suggest soft and friendly design is preferred to a harsher robotic aesthetic, with the idea that a soft shelled anthropomorphic or biomorphic design could make the robot more readily acceptable, although it should be noted that this is not from the end-user’s perspective. [25]

As mentioned above, Kimura et al. makes use of cat-like robot. The use of synthetic fur and feline appearance has been used to make it feel more natural to pet, stroke and hug it, as you would a real cat, which an example of attempting to improve acceptability through softer design (mentioned by Bradwell et al. [25] above). Some issues were noted however with children being fearful because of its realistic appearance. [1]

Ollie the otter is a furry otter-esque shape approximately the size of a baby, which the reason behind animal selection given as otters are “cute and familiar” but importantly generally not enough is known about them to point out unnatural aspects of Ollie’s behaviour, in contrast to if it was a cat or dog [26]. This is due to the uncanny valley effect [27] which occurs as a robot approaches near perfect realism, causing the robot to become disturbing, difficult to accept and uncomfortable to be around – a very relevant issue when considering use among children with ASD who may already struggle with social interactions [13]

There were several bird-type forms found in literature including PABI who has a penguin form [20] and KiliRo who is a therapeutic robot inspired by a parrot. KiliRo has a solid frame made from 3D printed polylactic acid and thermoplastic polyurethane and is comprised of a head, two eyes, beak, two wings, tail and two legs, split into three sections: upper, middle and lower [3].

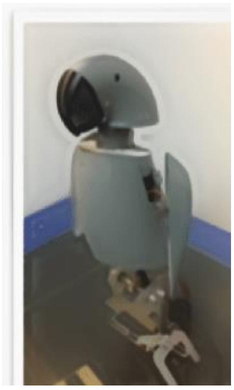
Other designs include dinosaur Pleo [15], robotic dog Aibo [14], teddy bear Huggable [28], raccon/gremlin style creature Leonardo [29], small cartoonish blob style Keepon [16], and elephant/creature Probo [30].









From left to right

Top row: AIBO [14], Huggable [28], Pleo [15]

Second Row: PABI [20], Leonardo [29], Ollie [26]

Third Row: Probo [30], Paro [9], Keepon [12]

Last Row: KiliRo [3]

The majority of animal robots feature a soft fur covering which can easily become a health hazard due to the collection of dust and dirt and therefore requires design measure to keep safe and clean such as the outer skin being removeable and washable [31], a point which also considered by Ollie, who’s fur is removable and who’s underneath is waterproof to protect from spill damage [26]. Arguments to include a soft outer at all include making the creature more soft and huggable, encouraging touch [32], making it more inviting to pet and stroke [1] and increasing comfort [31].

An interesting point which had not previously come to mind when considering design but was mentioned in several studies is the presented gender of the creature.. Sartorato et al. note that both the gender identity of the user and the perceived gender of the robot are important factors to consider to best facilitate interactions and maximise positives effects in individualised social therapy where children with ASD are involved, due to the high cross-over between ASD gender dysphoria [13]. To avoid stereotypes and allow users to assign a gender of their own choosing an androgynous design is suggested as preferable [25]

Features

The main features associated with the robotic pets discussed above are covered below, separated into sections covering movement, sensors, behaviour, niche additional features, and computing power.

Movement

Most robotic pets considered above featured some mechanical movement. Robotic pets with human abilities like forming facial expressions and moving body parts help to increase interaction and engage attention, eliciting greater responses in children with ASD than prompting by humans alone [33]. Some examples of interactive features include motions like walking, jumping, or dancing; body language such as tilting, turning, or shaking head; shrugging shoulders; and facial expressions comprised of smiling, frowning, or movement of the lips, eyebrows, eyelids, or ears [13].

Some issues can arise however when excessive movement occurs. For example, the robot Infanoid had a very high degree of complexity with 29 actuators allowing extremely wide range of movement and expressions which led to an uncanny valley effect [27] and caused overstimulation in children with ASD, leading to anxiety and embarrassment being experienced by the end user [16] [34]. Another issue with incorporating lots of movement into robots is the resulting hard mechanical structures which make up the body which can cause discomfort during long embraces and make it awkward to manually position limbs and parts into the position the end user desires [31].

Some examples of robots capable of movement include Pabi who has a variety of actuators which enable 8 degress of freedom, including movement of the eyes, beak head and wings. It uses movement of these parts to express its emotions rather than through facial expressions. [20]. KiliRo, the other bird type mentioned, previously has one degree of freedom in the head for left and right motion, one degree of freedom in the beak and each wing to enable up and down motions, and two degree of freedom in the legs for turning left and right and moving backwards and forwards. The motion was achieved via DC motors powered by a 2000-mAh lithium polymer battery. [5] Pleo is also capable of walking, as well as moving its tail and head, and blinking and moving its mouth [13]. Probo has one of the highest degrees of freedom, but is also one of the larger robots. It has a fully actuated head [32] capable of directing gaze, eye blinks, ear flapping, mouth movement and head nodding and shaking to form a variety of facial expressions [35], as well as a movable trunk [32].

Locomotion was not typically employed when it came to robotic animals and was featured infrequently, with examples including KiliRo [5], Pleo [13] and Aibo [29]. Issues surrounding mobility include difficulty navigating uneven floors and obstacles and present additional limitations which are absent from immobile robots like Paro [25].

Sensors

In order to have an affective robot pet that can interact with its surroundings and respond to stimuli, sensors are crucial and can be used to detect touch, motion, sound and visual inputs. This allows the robot to understand how the users are interacting with it and respond with appropriate movement and sounds [26].Many, such as Aibo [29], Pleo [29], Huggable [36] and KiliRo [5] include touch sensors, motion sensors, cameras, and microphones.

Whilst facial recognition is typically done successfully, it is noted by Bradwell et al. that voice recognition capabilities present in on-the-market solutions are limited and improvements are necessary for successful implementation, such that current voice recognition systems are more of a hinderance [25].

Another area of affective computing that requires development is embrace recognition and interpreting the meaning behind a specific gesture. Whilst most of the robotic pets mentioned contain some type of touch sensor, generally the focus is on acknowledging contact regardless of intent rather than identifying the specific gesture that occurred and its nuances. However, touch input can convey a significant amount of meaning, with studies showing that humans demonstrate emotional state through physical gestures [37] meaning robotic animals with gesture recognition capabilities have the potential to be a non-invasive physiological monitor for stress among paediatric patients rather than having biometric sensors and act as a "window into [a person's] emotional state" [17].

In order to be effective at identifying gestures, a robot must have suitable touch sensors. Robots such as PARO [9] and Aibo [14] use Force Sensitive Resistors (FSRs) which are inexpensive but poor at sensing lighter touches and differentiating gestures of similar pressure [17] which impacts the detail of gesture the robot can pick up on and is likely why PARO touch classification is limited to only 2 categories (stroke or hit) and Aibo's touch interaction is based purely on location of contact [38]. Other robots like Huggable [39] have much more advanced gesture recognition capabilities, in thanks to its 1500 sensors including temperature and capacitive sensors in conjunction with FSRs; however, this makes it very expensive and does not fall in line with the objective of this project. Several alternative approaches can be found within literature. One such approach by A.Flagg and K.MacLean [17] includes using a conductive fur sensor and a piezoresistive fabric pressure sensor, which changes in resistivity when pressure is applied and is cheap and flexible [40] making it a low-cost solution suited to soft toy design. Another approach used by K.Altun and K.E.MacLean [38] includes a combination of FSRs and an accelerometer. Lastly, Cooney et al. describe using sensors which measure distance rather than force or velocity via a “photo-interrupter emitting light onto, and receiving light from, the reflective inner surface of a soft outer cover” [30]. The approach with the highest reported accuracy comes from A.Flagg and K.MacLean [17] thus using the piezoresistive fabric pressure sensor seems preferable. The study also employs standard sequence statistics in contrast to other studies with rely on the relationships between multiple sensors to define gestures. This difference is hypothesised to be the “key to performance” [17] and could provide high results with much lower costing hardware than projects like Huggable [39]

Positioning of the sensors is of argued importance. Altun and MacLean disregard location information (meaning the same pressure and motion occurring at different locations will be categorised as the same gestures), arguing touch location between humans and pets does not convey emotion is the same way as between humans and humans and is thus less important. Additionally, they add that inclusion of location didn’t improve their personal results. [38] In contrast, Kleawsirikul et al. suggest that the loss of information makes it difficult to distinguish between types of embraces and it is therefore important to include location [31]. Actual sensor position varies between robots. Pleo features them on the head, chin, cheeks, shoulder back and thigh [29], Aibo’s are positioned on the back head and jaw [29] , whilst KiliRo’s are placed on the wings, forehead and body section [5]. Head and back placement seem to be the most frequent, and Yohanan and Maclean suggest sensor density should be focused on the back due to the frequency of touch and gestures occurring in this spot [37].

As well as touch sensors, some robotic animals also include motion sensors: Aibo includes two 6-axis motion sensor (3-axis gyro and 3-axis accelerometer) in the head and torso [29], Pleo has an orientation and motion sensor [29] and Yohanan and MacLean’s Haptic Creature includes a 3-axis accelerometer, with movement sensors being described as “crucial” for identify specific gestures like lifting and cradling, both of which convey emotion. [37]

Behaviour

The behaviour of robotic pets is typically of the form that the robot should react happy when receiving positive interactions from a user and sad when receiving negative input, such as aggression. An advanced robot Paro [9] is described as having a memory such that it can gradually adapt its behaviour, repeating movements and sounds that gain positive attention whilst trying to avoid pattens that elicit negative responses. Additional behaviour models are discussed below, presented in a study by M. Cooney, S. Nishio, and H. Ishiguro [30] detailing how different behavioural approaches of a social robot affects the relationship with the patient.

The three discussed approaches are: ‘Always affectionate’, ‘Tit-for-Tat’, and ‘Shy and hard to get’. The response of the study was that participants felt that ‘Always affectionate’ could become insincere and repeatable, whilst ‘Tit-for-Tat’ could be very volatile and presented a fickle character, destroying previous rapport. ‘Shy and hard to get’ could also be perceived as the robot disliking the participant when they didn’t understand that the robots unaffectionate behaviours were in jest. The key learnings from the study suggest that for the best rapport to be built between the robot and the participant the robot should convey genuine sincerity, variety and stability, with stability referring to the robot remembering previous interactions and not basing its response off a singular interaction. For example, the robot keeps a positive attitude but shows pain and sadness in response to a negative interaction preceded by positive interactions. It is noted that the robot should not be regarded as heartless or boring. There is also a disclaimer about the difference between western culture and eastern culture, since the participants were of eastern background and may have a different interpretation of appropriate behaviour compared to users of a different culture.

Additional Features

Additional features typically included vocalisation and noises by means of a speaker. Most robotic animals feature only noises and sounds (such as Pleo who pseudo-vocalised by using “Hee!” to expressive a positive emotion and “unh uhn” for “no” [13]) however some can talk such as KiliRo who has a text to speech module [5] and Probo who can deliver pre-recorded phrases in a neutral male voice [35]. In general, speech varied from robotic to more human-like, displaying varying levels of emotional prosody [13].

Autonomous charging and “standby” modes are additional useful features noted by Bradwell et al., who noted that battery life is a concern, given that it could cause distress if a robotic animal “died” whilst in use [25]. Some robots such as Aibo already perform self-charging when on low battery [29].

A lesser occurring feature was the ability to transmit temperature, using temperature as a way to convey emotion. For example, the robot becomes warmer when it’s happy, making it more friendly and comforting to hold and hug. [30]

Computing power

Raspberry Pi was featured commonly, with both Ollie [26] and KiliRo [5] using one. KiliRo also features two Arduino family of microcontrollers [5]. A more powerful robot, Huggable features an embedded PC with 802.11g wireless networking [36]. Pleo also has its own software platform. [29]

Software and Hardware Design Research

Software

Analytical methods

Methods discussed for obtaining a reliable identification system for facial expression and embrace recognition included neural networks, random forests and K-means clustering. Some learning was done in a supervised way whilst other methods were unsupervised.

There are currently various models representing known emotions and their transient in time. One idea discussed was emotional space, wherein emotional states can be organised according to their distance from each other, allowing them to be represented in their continuous dependent nature. [41]

Data Sets

In order to train an identification system for the recognition of embraces, facial, voice or otherwise, a data set is required. Of the studies covered above, many created their own data sets from willing participants and trials; however this is not likely to be viable given the time constraints of the project and the ethical approval required to run such a set up. There are existing datasets for use of this sort of project, namely facial recognition datasets. This includes AffectNet which is a database covering a vast range of facial expressions in the wild. It’s available on request however there is a delay in availability due to the demand for the data [42]. Another database is Fer-2013 which consists of over 34000 images categorised into 7 emotions (happy, sad, angry, afraid, surprise, disgust, and neutral) and is available publicly for free and with immediate access. The database been organised automatically, so the labels are not guaranteed to be accurate. Images have also not been purposefully created but rather gathered from multiple sources so may not be of high quality. The images are also small (48 x 48) and in greyscale. [43] An updated label set for the dataset exists: Fer-2013 + which relabels the database using 10 crowd-source taggers, providing a better-quality match for the images compared to the original labels; however this does require additional processing to integrate the new labels with the existing database. [44]

Hardware

This section takes a brief look into the available hardware.

Sensors

Most projects a wireless camera or a USB camera incorporated as well as a microphone for picking up visual and audio inputs. Piezoresistive fabric along with FSRs are likely to be best touch sensors for this application

Speaker

Most robotic animals discussed contain a speaker to emit emotions. Microcontrollers like Raspberry Pi can be used for audio control purposes.

Eyes

Options for eyes include mechanical eyes as seen on robots like Paro [9] and Furbys [45] or LED eyes such as those used by Aibo [14].

Mechanical eyes require actuators and motors to be expressive which can add up in cost and take up room within the head of the animal. The design process to make the eyes could also be lengthy. This could be circumvented by using already existing designs, for example using 3D printer design templates.

LED eyes can be very expressive without the need for actuators – just a screen and a microcontroller such a Raspberry Pi or an Arduino. This has the disadvantage of being more susceptible to damage compared to its mechanical counterpart.

Programming scripts for the expression of both mechanical and LED eyes are wildly available.

Battery, Motors and Actuators

There is little discussion of battery and power sources surrounding the robots mentioned above. It is likely that for data to be transmitted back for analysis the robotic animal will be attached to an external power source, hence battery will not be considered for this project, but is something to be looked into in order to make the design more mobile.

Similarly, motors and actuators forming the mechanical structure of the design are not the main focus of the project and have not been given research time due to constraints.

'Skin' and 'Skeleton'

The robot body will have to feature an outer skin. To encourage cuddling and holding, it will be of synthetic fur.

The skeleton of the animal could be either plastic or metal, however metal would be heavy and expensive, whereas using plastic allows for 3D printing, is cheaper and is lighter.

Chapter 3

Method

Due to time constraints on the project, the focus was placed on developing a facial expression recognition software and producing a physical design concept. This allows a simulation to be run depicting how the physical design would respond in reality if hardware had been able to be implemented and gives direction and a base for future work. The software and data processing required were decided to be the lengthiest part of the design process so was tackled first using the process described below

Software development

MATLAB 2020b was chosen to for the data processing and network learning of the software section of this project due to knowledge and experience with the program compared with other methods, such as python. MATLAB also contains a variety of pre-trained easily accessible neural networks, ideally suited for image classification.

Data Processing

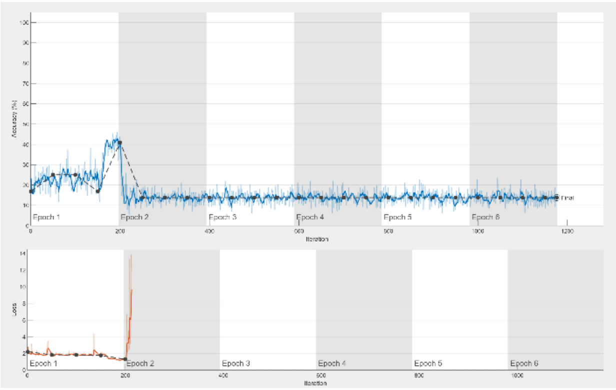
The FER-2013 dataset was chosen due to it being publicly accessible and immediately available. The new FER+ labels were not implemented to cut down on pre-processing time, since the dataset already required a significant amount of processing to be compatible with deep neural network training. Firstly, the csv file was loaded into MATLAB and re-structured into a new table with each pixel and emotion label clearly organised into separate columns. A dummy pixel of value 0 was also added as the first pixel to bring the number up to a square of 48. Each row of the table (where one row represents one image) was then reshaped into a 48 x 48 matrix and stored in a cell array using MATLAB's inbuilt reshape function and for loops. The data was processed in phases of approximately 5000 rows/images due to hardware limitations of the computing device used to process the data, which had insufficient memory to handle all the data at once. Each cell array was then collated together in a singular cell array entitled imagestotal. The find function was then used to collect the indices relating which images correlated to each emotional label. This was then used to separate imagestotal in 7 cell arrays entitled emotion0, emotion1, emotion2, emotion3, emotion4, emotion5 and emotion6, each containing their respective images in matrix form. Each cell array was then processed using the imshow function to convert each matrix into a visible image before using sprintf and saveas to save the image into its respective emotional category folder as a png file with a title of the format emotionm_n where m represents the emotion category and n represents the index of the image. The end of the process gave 35887 png format images organised into emotional category folders.

The data was then loaded into an image date store, using the folder names as labels. The splitlabel function split the data into a training set and a validation set with a 70/30 split respectively. The augmented data store was then used to pre-process the images, since all pre-trained networks have an input of size m x n x 3 (representing an RGB image) and the images contained within FER-2013 are grey scale and hence of input size m x n x 1. The 'ColorPreprocessing', 'gray2rgb' functionality of the augmented data store was used to rectify this. At this point the data was ready to be imported into the deepNetworkDesigner app. The MATLAB script containing the instructions for the above processes can be found in the appendix section under ‘dataProcessing.m’.

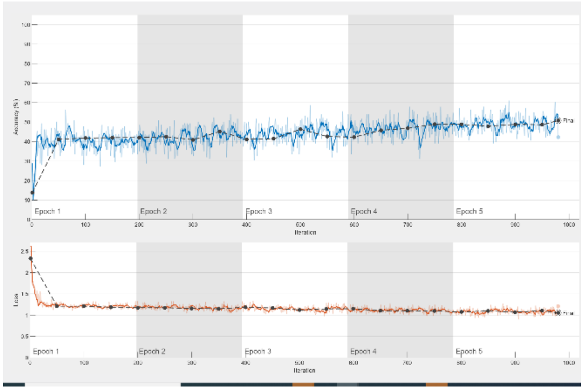
Network Training

The deepNetworkDesigner app was used to train the emotion recognition network by using image classification. Transfer learning was chosen as it is faster than training a network from scratch and only required some modification of existing networks, and MATLAB gives access to multiple pre-trained networks. SqueezeNet and GoogLeNet were picked to run experiment trials with. SqueezeNet was typically less accurate but faster than GoogLeNet. Due to the difference in input size (227 x 227 x 3) for SqueezeNet and (224 x 224 x3) for GoogLeNet, the initial input layer of both networks were replaced to expect an input of (48 x 48 x 3). The final classification layer was also replaced so that the classification labels matched that of the emotion categories. GoogLeNet required replacement of the fully connected layer such that it classified the information into 7 categories instead of 1000. SqueezeNet required replacement of a filter layer, again to reduce class size from 1000 to 7. Multiple training options were also tested, including solver (sgdm, adam and rmsprops) , minibatch size (10, 128, 256) and learning rates (0.01 – 0.0001) [46]. Mini training runs of up to 6 epochs were used to efficiently evaluate the best options as network training is a time-consuming process. Once the best options were identified the network was allowed to train for longer durations of 20-50 epochs. Overtraining issues occurred once training durations exceeded 15 epochs so longer network training began to be capped at 15 [47]. Due to unsatisfactory accuracy results in the network results, the number of classes was dropped from 7 to 4 with emotions Surprised, Disgusted, Neutral being deemed of lesser value in the context of robot assisted therapy and thus removed from the data set, leaving Angry, Sad, Happy and Frightened. This increased the base accuracy from 14.3% to 25% . The image data store was re-done so only the 4 desired categories of emotion images were imported in the deepNetworkDesigner app and the above process detailing the evaluation of the training options was repeated for the new 4 class network until a net of satisfactory accuracy was achieved. The network as the exported into the workspace to be used to classify images in a MATLAB.

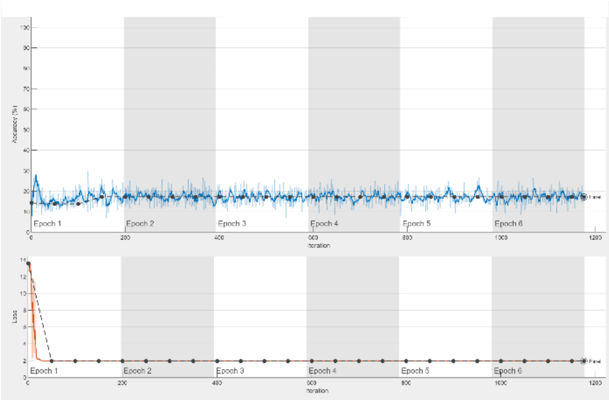
The following figures summarise the process, giving an insight into some of the best and worst trials runs used to narrow in to an optimal final solution.



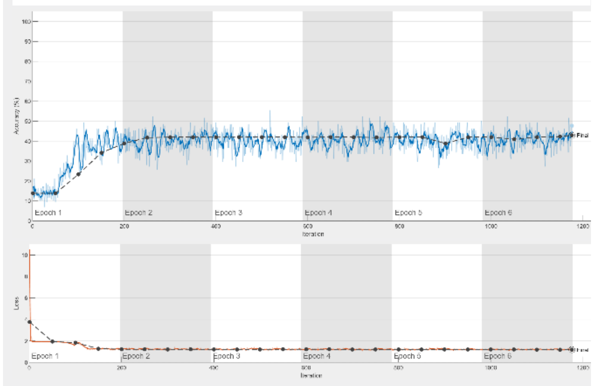
GoogLeNet, 7 classes, with the following training options: solver - sgdm, learning rate - 0.03, minibatch size – 128 Final accuracy of 13.80%



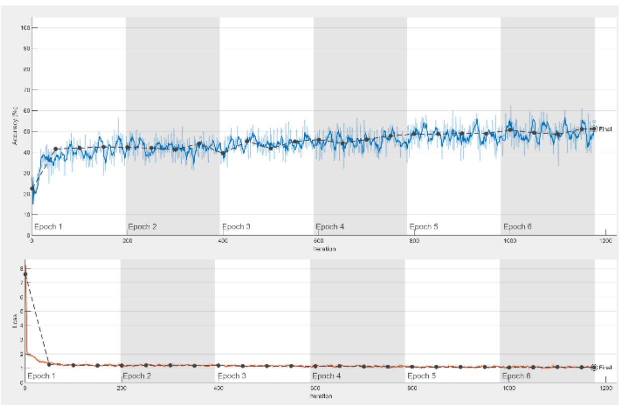
GoogLeNet, 7 classes, with the following training options: solver - sgdm, learning rate - 0.005, minibatch size – 128 Final accuracy of 50.77%



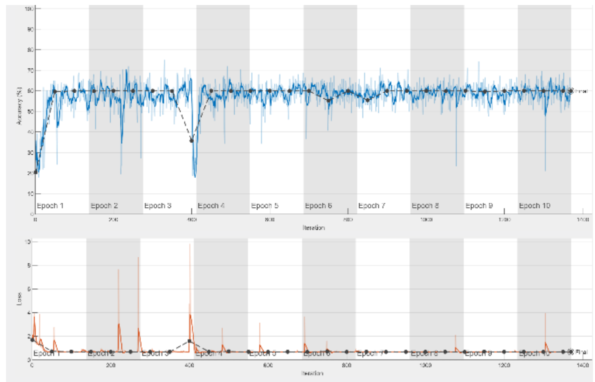
SqueezeNet, 7 classes, with the following training options: solver - sgdm, learning rate - 0.03, minibatch size – 128 Final accuracy of 17.27%



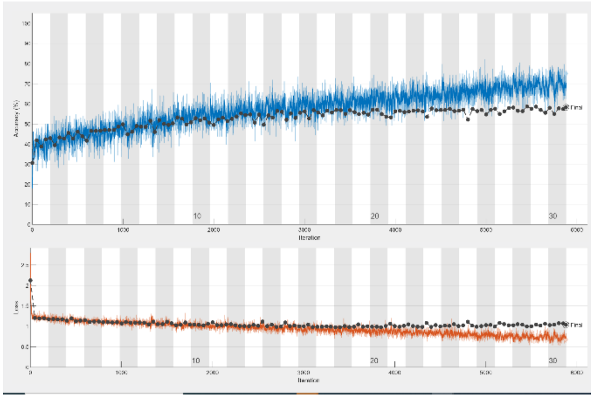
SqueezeNet, 7 classes, with the following training options: solver - sgdm, learning rate - 0.01, minibatch size – 128 Final accuracy of 42.93%



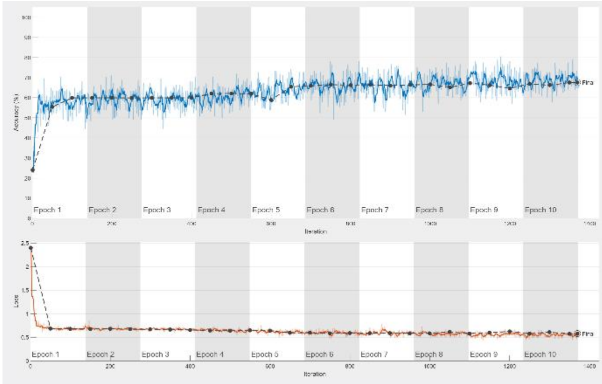
SqueezeNet, 7 classes, with the following training options: solver - adam, learning rate - 0.001, minibatch size – 128 Final accuracy of 51.19%



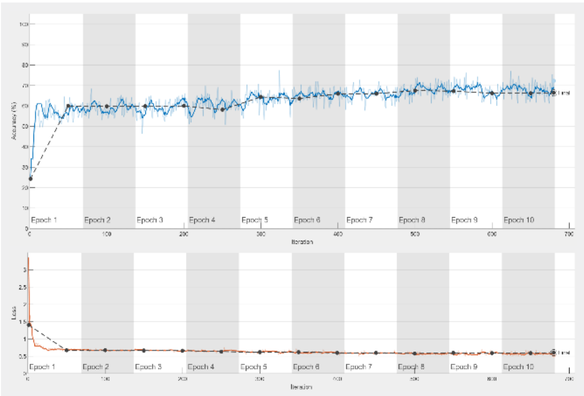
GoogLeNet 4 class with the following training options: solver - rmnsprops, learning rate - 0.001, minibatch size – 128 Final accuracy of 59.92%



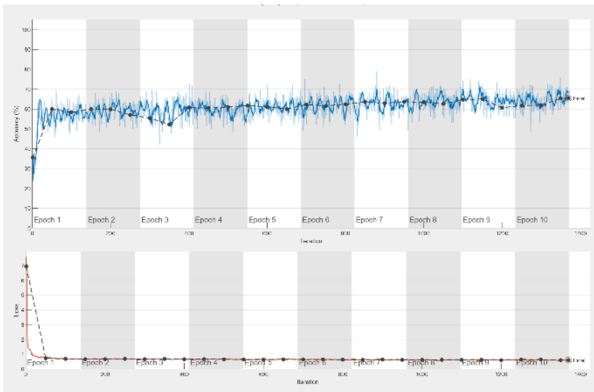
GoogLeNet, 7 classes allowed to train over 30 epochs with the following training options: solver - sgdm, learning rate - 0.005, minibatch size – 128 Final accuracy of 58.27%



GoogLeNet, 4 classes, with the following training options: solver - adam, learning rate - 0.001, minibatch size – 128. Final accuracy of 67.57%



GoogLeNet, 4 classes ,with the following training options: solver - adam, learning rate - 0.001, minibatch size – 256 Final accuracy of 66.45%



SqueezeNett, 4 classes, with the following training options: solver - adam, learning rate - 0.001, minibatch size – 128 Final accuracy of 65.58%

Design concept process

It was acknowledged that given time constraints there would be no time for hardware implementation however the aims of the project could still be met if a valid concept could be produced, providing a

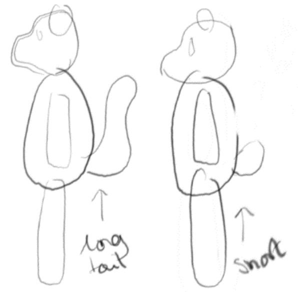
solid foundation for future work. The main capabilities to be considered was the focus on embrace recognition and the hardware required for successful data collection to enable this.

Using the information from the literature review, the key features chosen to be included in the concept were body shape and form, movement capabilities (including expressions), sensor type, audio capabilities, computational power and desired behaviour.

Body form

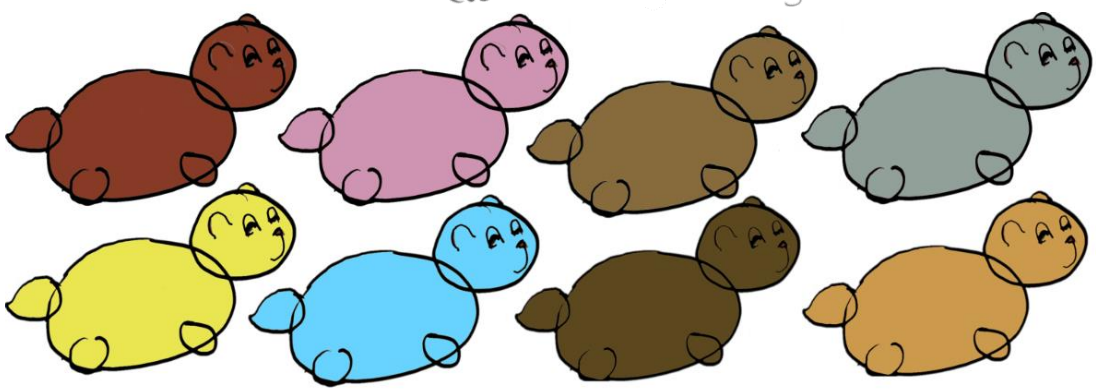
The physical body of the pet was considered first and involved several concept sketches. drawing inspiration from traditional children's soft toys and also current robots like Paro [9], Huggable [28], Ollie [26] and the Haptic Creature [37]. A non-domesticated creature was deemed most appropriate, following in Ollie and Paro's steps to gives something that is familiar and cute but far enough removed from common knowledge to prevent uncanny valley [26]. The body should be as soft as possible for comfortable cuddling and hugging over a long duration, with as few hard moveable mechanisms as possible, for comfort purposes as noted by N. Kleawsirikul; H. Mitake and S. Hasegawa [31]. Silicone limbs like Ollie [26] shown in one concept allows for sensors in the arms as well as the main body, proving more gesture data but being more conducive for cuddling than hard solid limbs like Huggable [28]. Taking away the limbs entirely to resemble a more simplistic design like Keepon [12] removes this issue and simplifies the design but also reduces possible gestures. However, following Yohan and Maclean's observation that majority of touch occurs along the back, this may not as big of an issue [37]. A furry outer "skin" invites stroking and touch but should be easily removable to be able to washed. Different colour schemes were considered as shown. Other animal features including ear designs and tails were also considered.

Body type sketches showing different considerations of limb length and body shape



Side profile and tail concept drawings

Different ear concepts



Colour experimentation

Movement capabilities

To provide as an expressive face as possible, the minimum amount of movement required by the face included moveable eyebrows, eyes, ears and mouth. LEDs eyes were chosen like those of Huggable [28] and Aibo [14] to give it emotionally expressive eyes without the need for complicated mechanical infrastructure, increasing comfort and reducing cost. There is lots of information available to implement this idea, including programming guides [48] and easily purchasable hardware, making it a realisable idea for future. Other facial movement (that of the eyebrows, ears and mouth) will be via mechanical means such that the face of the creature doesn't become too robotic. This will require the "skull" of the animal to be capable of holding actuators and the mechanical structures required; however, the technical specification for this will not be covered by this project.

The face needs to be able to be manipulated into expressing expressions including happy, sad, blinking, and resting. Some concepts sketches for the different expressions are show below

Different facial concepts shown from left to right Happy, Sad, Resting

Sensor type

The touch sensor selected was a piezoelectric fabric sensor which can be easily wrapped around and sewn. A motion sensor has decided to be included following recommendations from Yohan and Maclean, giving the robot the ability to sense movement and provide data for recognising lifting and cradling [37].

Other sensors include a camera fitted to allow expression recognition. Possible locations considered included the eyes and nose. A microphone could be added if voice recognition and tone recognition were desired in future software advancement.

Audio Capabilities

A speaker has been included to allow the robot to express itself audibly. Sounds considered to express sadness included: whimpering, crying and yelping. Sounds considered for happiness included: purring, laughing, and joyful exclaiming. Several sites were found containing audio files in the public domain and/or licensed for commercial and educational purposes including mixkit.co [49] and BigSoundBank.com [50]

Computational power

All data processing and networks have been trained through MATLAB when considering the software phase of this project, hence it would make sense to carry this across to the implementation of the design. Certainly MATLAB is a very powerful program that is capable of the task; however other computing options like Raspberry Pi and Arduino would allow the design to become more mobile and independent. In order to use the robot as a non-invasive physiological sensor however the data needs to be outputted to an external platform for reading by, for example, a therapist or medical practitioner. In which case, MATLAB may as well be implemented for data collection purposes. However, control of the robot's responses (for example the control of the actuators in the face) could be delegated to an inbuilt computing source, like Raspberry Pi or Arduino microcontrollers. Technical details of how the hardware could be implemented is beyond the scope of this project and is up for future consideration.

Desired behaviour

The behaviour should be act as a display of empathy and show a level of interaction with the user. Stability and sincerity are the focuses following review by M. Cooney, S. Nishio, and H. Ishiguro [30], to build a positive rapport with the end user. A typical structure, taking influence from available robots, would involve remaining generally positive but expressing sadness at negative interactions (for example aggressive contact or recognition of a sad/angry expression) and increased joy under positive conditions (such as stroking, gentle petting and recognition of happy faces).

Digital Simulation

Once the design concept was solidified, attention was turned to producing a simulation of the desired responses and facial expressions from the pet. This is to showcase the design concept and expression recognition network and give an idea of the intended implementation. Each finalised expression was deconstructed into animation frames, drawn using SketchBook [51], exported as png files and imported into MATLAB where they were developed into functions to display them as animated sequences in a window figure (the sequences can be found under the Results section of this report). The finalised audio files were also imported into MATLAB and included as part of the function. A script was then developed which enabled the webcam to gather images to apply the facial expression recognition network to in real time and run the animated sequences as appropriate. The script detailing this process can be found in the appendix labelled ‘liveActionResponse.m’.

Ideally this simulation would have been presented to a group of voluntary participants in order to test the emotion recognition software in a real life scenario and also gauge responses to the animated sequences; however ethical approval would needed in order to do this, which was not sought for this project due to time constraints.

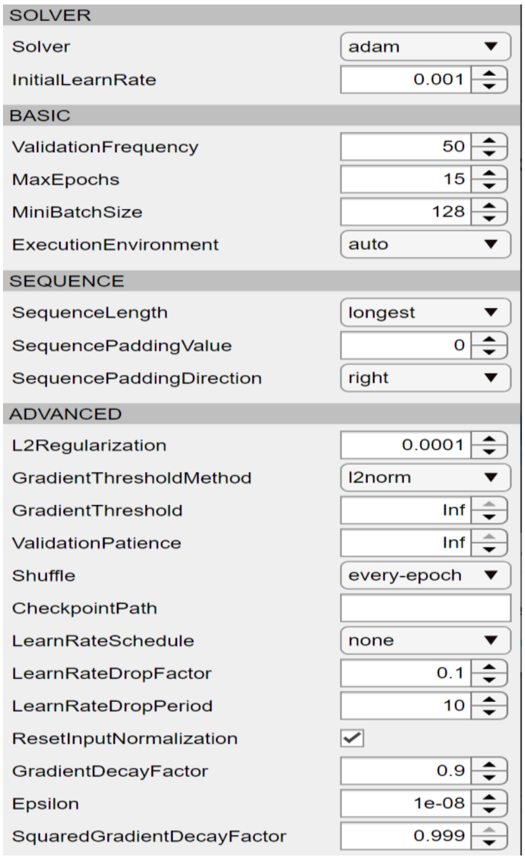
Chapter 4

Results

This section provides a concise representation of the finalised details from the project, including the final approved network, final design concept and hardware specification where appropriate, and the animated sequences implemented in the digital simulation.

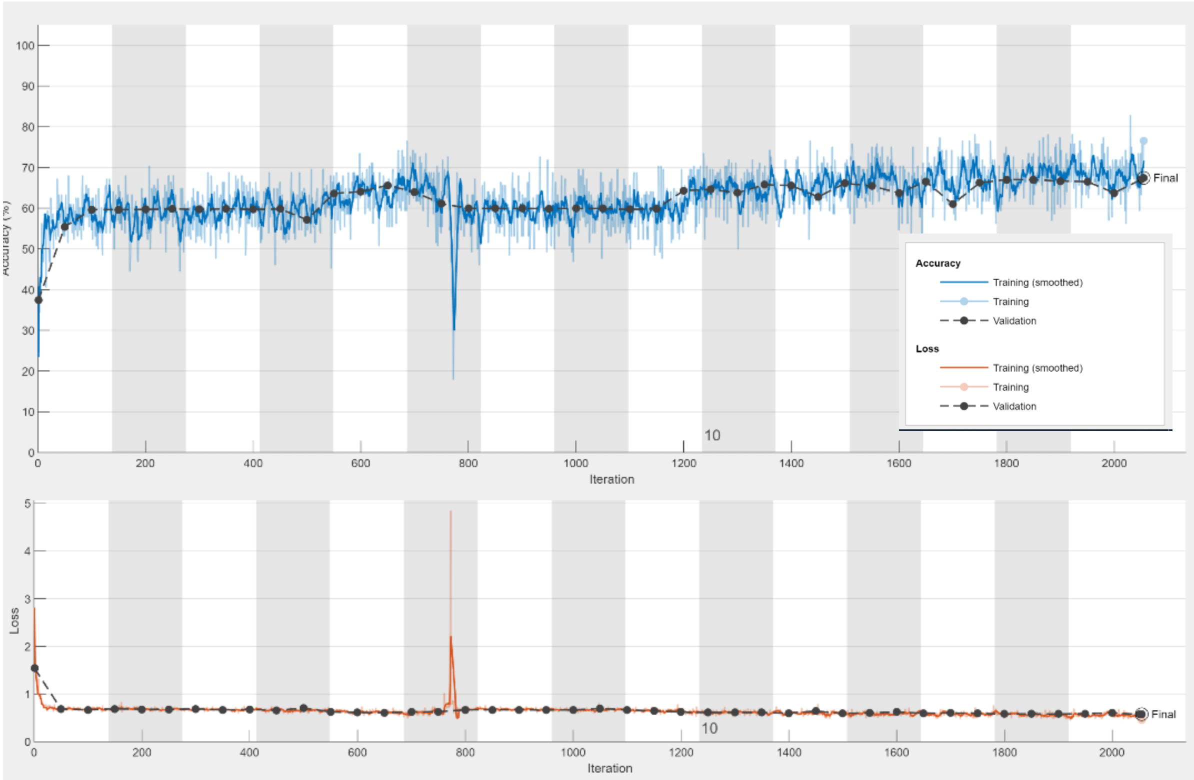
Network

The final trained network consisted of a modified version GoogLeNet trained using the following training options:



Training options for final network

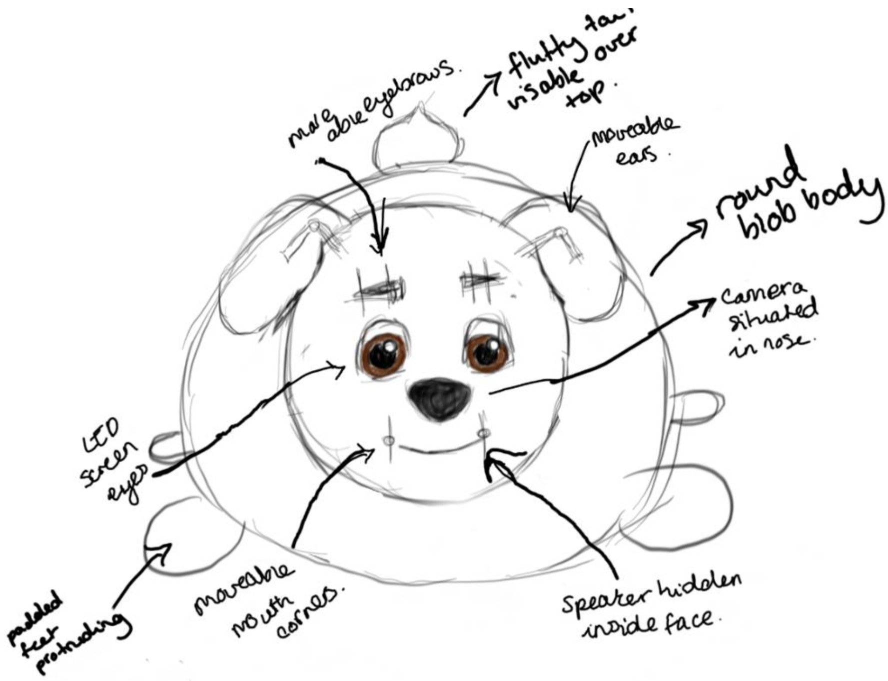
The network achieved a validation accuracy of 67.50% compared to a baseline accuracy of 25% and was trained over the course of 15 epochs for a duration of 169 minutes and 9 seconds, completing 2055 iterations, with 137 iterations occurring per epoch.



Final Trained Network

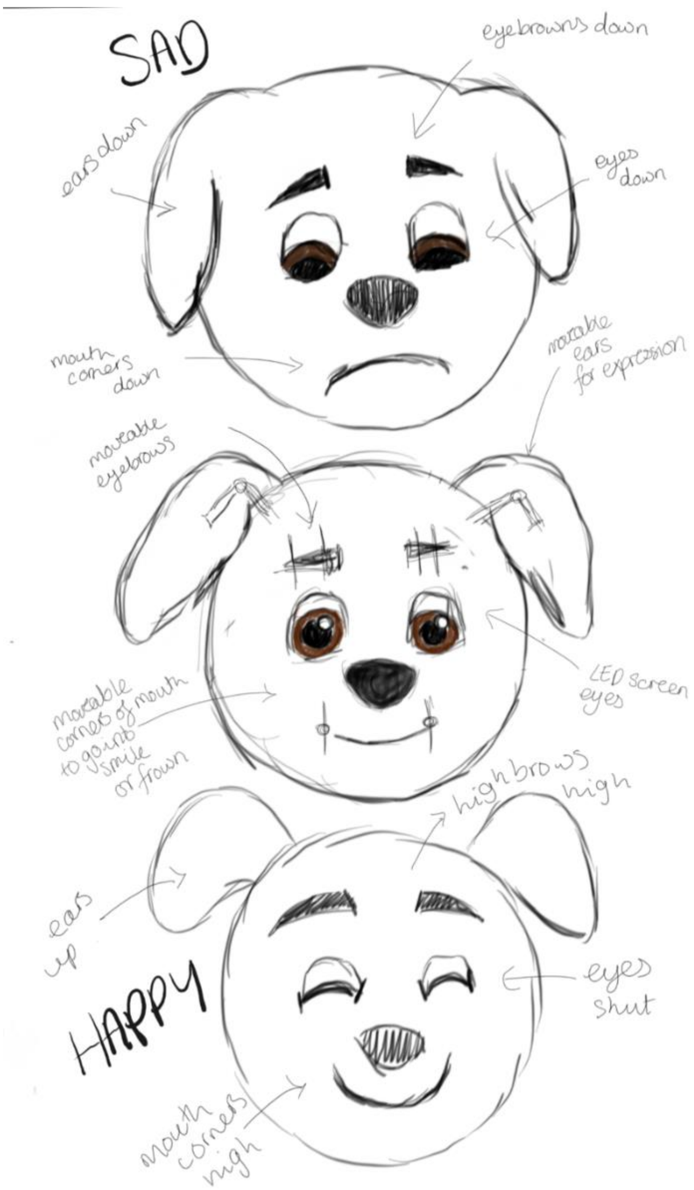
Final Design Concept

Final design concept showing side and aerial view



Final design concept showing front on view

The final design was chosen to be a rounded barrel shaped guinea pig-esque body with ears loosely resembling those of a dog, and 4 small, padded feet, with no true limbs. The colour chosen was a chocolate brown, with a synthetic fur outer. Eyes are represented in the form of an LED screen. A camera is situated in the nose and a speaker underneath the chin but no microphone. Piezoresistive fabric wrapped about the main body and situated on the head acts as the main touch sensor, with a 3D accelerometer motion sensor also located in the body. Actuators for the facial movement are controlled by a Raspberry Pi. The eyebrows, corners of mouth and ears are capable of movement. Ports are situated near the rear of the body for the output and input of data. The audio file chosen to express joy is a cat purring [52] and the audio file to express sadness is a dog whimpering [53]. The final concept sketches for facial expressions are shown below.



Final concept sketches of facial expressions

Animation Sequences

The following sequences are comprised of frames representing the deconstructed version of the final expression concepts and are representative of the expression hoped to be achieved by hardware if the project were continued.

Resting (open eyes then blink with a small smile)

Smiling characterised by upwards closed eyes, higher eyebrows, and a higher smile

Sad characterised by lowered eyebrows, downwards facing eyes and a frown

Chapter 5

Results Analysis

Software

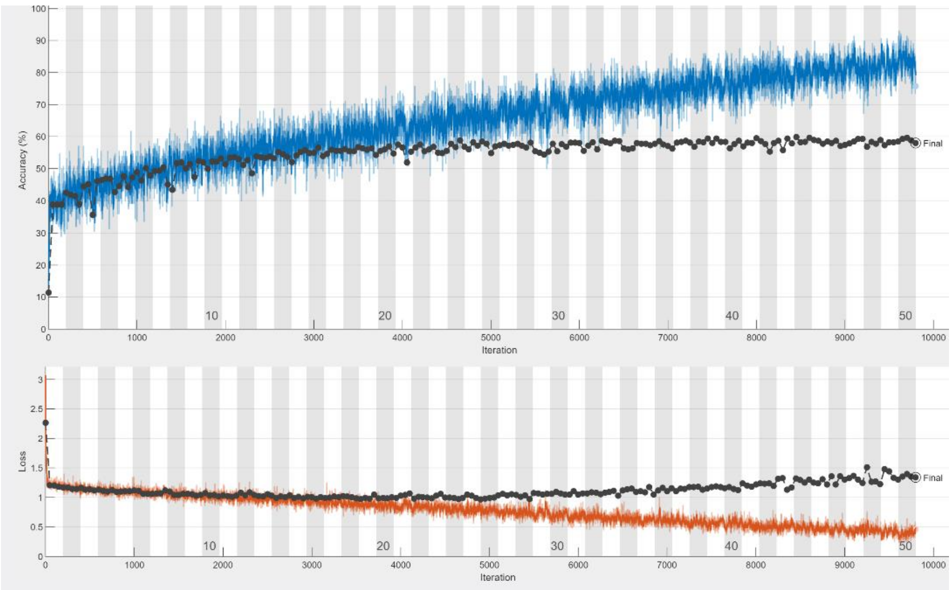
The software produced in the project was below the standard desired at the start of the project. Ideally an accuracy of at least 80% would have been acquired compared to the achieved 67.5%. This could have been due to the vast difference between the expected input to the network (as it was pre-trained) and the actual input. Notably, there was a large pixel size difference (224 x 224 vs 48 x 48) as well as the images from the dataset being greyscale compared to RBG images. Whilst the augmented data store was used to process the images to counteract this, it's unlikely to be as effective as using RBG images from the outset. These differences are likely to affect the quality and efficiency of transfer learning, and it's likely that better results would have come from building the network from scratch. However, this would have been a very time-consuming process to fit into an already very tightly constrained project and could have had knock negative impacts on the remainder of the project.

Another key factor which is likely to have had a detrimental impact on the network is the original labelling accuracy of the data set, and the quality of the images. Close inspection of the database shows some labelling which could be considered incorrect. Examples of this are shown in the figure below

Images from the FER-2013 database with original labels shown at the top and alternative "corrected" labels from FER+ shown below [44]

FER+ [44] is an improvement on the original labels and could have enhanced the dataset if utilised but would have added processing time to the data to integrate the new labels with the old dataset.

Another issue identified during the training process is the tendency of the model to overfit the data when given long periods to train. The training accuracy data starts to overtake and soon greatly exceed the accuracy of the validation data, meaning the accuracy of the networks couldn't be improved by training for longer durations. This is demonstrated in the figure below.



Example of an overtrained model

It can be seen that the validation accuracy plateaus at approximately 55% around 15 epochs in whereas the training accuracy continues on an upward trend. This can be caused by the model being too flexible for the data and starting to overfit the training data taking away from its ability to define a generalised solution which can be applied to a validation set [47]. This is likely given GoogLeNet is intended to categorise into 1000 different classes as opposed to 7 or 4. This is another issue that could be addressed by building a network from scratch to achieve optimum flexibility.

Another point to note is the small pool of variations which were tested in the training approach. For example, only GoogLeNet and SqueezeNet were the only pre-trained networks to be experimented with, which, given the vast array of pre-trained networks available, is a very small sample. If more time had been dedicated to exploring existing networks, a more suitable alternative might have been found – for example, a network that uses much smaller input sizes or grayscale images.

The software also sacrificed some functionality to increase accuracy, as noted in previous sections, with certain emotions being excluded later in the design process. This is something which, had it been considered earlier in the process, could have reduced time spent processing images; however, it was expected at the time that the network would be capable of accurately sorting the images into their relevant emotions.

The final network was not able to be tested in a real-life scenario due to lack of ethical approval; hence observations about its accuracy are purely theoretical and limited comments can be made about its real-world success. Importantly however, a working network was acquired and the means to test it (the digital simulation) are ready to be utilised in future work. The software developed aligns with the objective to develop a facial expression recognition feature and hence is conducive with the aim of the project.

Design Concept

Design

The finalised concept produced aligns well with the aims and objectives of the project, particularly in providing a solid foundation for future work. It combines a vast array of knowledge garnered from the literature review and represents an amalgamation of features found in the best robotic pets currently available on the market.

Some flaws include the lack of detail pertaining to implementation of the concept, particularly actuator, mechanical and wiring details, as well as measurements. The technical knowledge is lacking here and is something that requires particular attention in future work; although it should be noted that the design has been conceived using a realistic approach and should not be overly difficult or complicated to achieve. If anything, the design is quite basic and more features could be added, such as more facial expressions, greater degrees of movement, and more hardware (such as a heating element and microphone). With regards to the behaviour of the robot, memory could be considered, and a system developed such that the robot is capable of remembering users and building greater rapport.

The design does lack validation from a user perspective, and it would have been preferable to survey people and gain opinions and views about the concept. This should be considered a priority for future work has the end user’s receival and acceptance of the design is key to the success of an affective pet like this.

Hardware

Specific hardware selection is lacking and requires more in-depth research. Not enough time was able to be dedicated in this project for a thorough review of all components. If components were in a position to be selected, they should be chosen mainly on a cost basis, to fulfil the aim of the project to produce an affordable solution. In this instance, performance was not the key, unlike with most experimental projects such as Huggable [28] and Leonardo [29] who represent extremely expensive high-end robotic animals. It is expected that the sacrifice in hardware performance can be made up with developments in the software and data processing of this robotic pet.

Discussion of project outcomes against aims and objectives

As discussed above, it is felt that the results achieved from the project satisfied some objectives of the project. The research phase was mostly comprehensive in relation to assessing current robotic animals available, their features, purpose and effectiveness however the time constraints clearly affected the software and hardware technicalities portion of the literature review, with in-depth analysis into analytical techniques and particular components missing. Overall, the research phase objective is considered to be complete.

The application of knowledge from the research phase with regards to the design concept has also been achieved, despite gaps in the literature review reflecting in the lack of technicalities within the design concept.

The software development phase of the project is the most lacking, with only facial expression recognition being approached despite multiple features being suggested for development. This was affected by the lack of data available due to not being able to obtain and use hardware to gain own personal data, and lacking the time to find datasets available for public use in relation to voices and touch sensor data for the purpose of developing voice recognition and touch recognition software respectively.

It is also noted that the outcomes are unable to be tested by the end user in this project due to a failure to seek and gain ethical approval, meaning the results are unable to be validated in a real-world context. As mentioned previously, this should take top priority for future work.

Chapter 6

Conclusions

Robotic therapy animals have some very useful practical applications in the real world and, despite being a recent application of affective computing (the majority of literature being under 20 years old at the time of writing), already have a wealth of research behind them, with some impressive projects like Huggable [28] and Paro [9] providing excellent examples of their capabilities. However, there still remains a vast amount of untapped potential to be exploited, particularly with regards to use of information transmitted by touch and tone of voice, and they are yet to feature in mainstream educational, medical and social care environments. This project provides a window into the current state of robotic animals and gives a foundation for future work. The scope of the project was severely reduced from the original outcomes set (notably the lack of implemented hardware) due to time constraints, however overall it succeeded in accomplishing several objectives. The results achieved with regards to software are not optimal but have scope for improvement and the design concept presents a genuine solution in the line with the aim of the project to provide an affordable alternative to more expensive social robots on the market. With further work it is believed the design will be fully capable of enhancing the learning environment and being a valuable tool for communication between children and medical practitioners.

Recommendations for further work

As noted throughout, the project has a lot of potential for improvement.

Software development is a key area that can be improved up. More experimentation needs to be carried out with network design, including researching different pre-trained networks and building networks from scratch. Additional analytical methods besides neural networks should also be explored such as K-means clustering. Effects of supervised versus unsupervised learning should also be explored. More time needs to be dedicated to the research of these methods and how best to implement them.

There are still a number of features from the objectives that did not feature in the project, including voice recognition and remote control capabilities, which are areas that could be researched and added in to the design to further add to the functionality of the robotic animal.

A lack of ethical approval hindered areas of the project that required data collection, meaning that features such as embrace recognition were unable to be developed, so this another avenue which could be explored in future work. Testing of the outcomes was also not able to be carried out due to this, so this is a priority in the future.

The design concept also has the scope to be filled out with more technical hardware details and ideas for implementation. A concrete design for the underlying ‘skeleton’ and construction of the body is yet to be made, and this would be required to bring the design into existence. A look into the mechanical components represents the next logical step for construction. A more detailed review of hardware available would enable this.

Overall the project has the scope to be taken further and this report serves as a solid foundation for future reference.

References

[1] R. Kimura, N. Abe, N. Matsumura, A. Horiguchi, T. Sasaki, T. Negishi, E. Ohkubo and M. Naganuma, “Trial of robot assisted activity using robotic pets in children hospital,” in *SICE 2004 Annual Conference*, Sapporo, Japan, 2004.

[2] C. J. Moerman, L. v. d. Heide and M. Heerink, “Social robots to support children’s well-being under medical treatment: A systematic state-of-the-art review,” *Journal of Child Health Care,* vol. 23, no. 4, pp. 596-612, 2019.

[3] J. Bharatharaj and e. al., “Social engagement of children with autism spectrum disorder in interaction with a parrot-inspired therapeutic robot,” *Procedia Computer Science,* vol. 133, pp. 368-376, 2018.

[4] H. Woo, G. K. LeTendre, T. Pham-Shouse and Y. Xiong, “The use of social robots in classrooms: A review of field-based studies,” *Educational Research Review,* vol. 33, no. June, 2021.

[5] J. Bharatharaj and e. al., “Sociopsychological and physiological effects of a robot-assisted therapy for children with autism,” *International Journal of Advanced Robotic Systems,* vol. 14, no. 5, 2017.

[6] B. Scasselati, H. Admoni and M. Mataric, “Robots for use in autism research,” *Annual review of biomedical engineering,* vol. 14, pp. 275-294, 2012.

[7] e. a. Syamimi Shamsuddin, “Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO,” in *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*, Malacca, Malaysia, 2012.

[8] L. Nasi and e. al., “Pomelo, a Collaborative Education Technology Interaction Robot,” in *14th ACM/IEEE International Conference on Human-Robot Interaction (HRI),* Daegu, Korea (South), 2019.

[9] IEEE, “ROBOTS: Paro,” 2021. [Online]. Available: https://robots.ieee.org/robots/paro/. [Accessed 18 October 2021].

[10] RoboPets, [Online]. Available: https://www.robopets.co.uk/product-page/silver-cat. [Accessed 18 October 2021].

[11] J. F. S.J. Whitcher, “Multidimensional reaction to therapeutic touch in a hospital setting,” *Journal of Personality and Social Psychology,* vol. 37, no. 1, pp. 87-96, 1979.

[12] IEEE, “ROBOTS: Keepon,” 2021. [Online]. Available: https://robots.ieee.org/robots/keepon/. [Accessed 18 October 2021].

[13] F. Sartorato, L. Przybylowskia and D. K.Sarko, “Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots,” *Journal of Psychiatric Research,* vol. 90, no. July, pp. 1-11, 2017.

[14] IEEE, “ROBOTS: AIBO,” 2021. [Online]. Available: https://robots.ieee.org/robots/aibo2018/. [Accessed 18 October 2021].

[15] IEEE, “ROBOTS: Pleo,” 2021. [Online]. Available: https://robots.ieee.org/robots/pleo/. [Accessed 18 October 2021].

[16] H. Kozima, M. Michalowski and C. Nakagawa, “Keepon: a playful robot for research, therapy, and entertainment,” *Int J of Soc Robotics,* vol. 1, no. 1, pp. 3-18, 2009.

[17] A. Flagg and K. Maclean, “Affective Touch Gesture Recognition for a Furry Zoomorphic Machine,” *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction,* pp. 25-32, 2013.

[18] Y. Zhang, F. Yan, S. Li, Y. Wang and Y. Ma, “Effectiveness of animal-assisted therapy on pain in children: A systematic review and meta-analysis,” *International Journal of Nursing Sciences,* vol. 8, no. 1, pp. 30-37, 2021.

[19] S. Y. a. R. Kimura, “Investigation of playing with entertainment robotic pet of pre-school aged child in nursery school by video observation,” *SICE Annual Conference 2007,* pp. 654-659, 2007.

[20] L. Dickstein-Fischer and G. Fischer, “Combining psychological and engineering approaches to utilizing social robots with children with Autism,” *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,*

EMBC 2014, 2014.

[21] Z. Salimi, E. Jenabi and S. Bashirian, “Are social robots ready yet to be used in care and therapy of autism spectrum disorder: A systematic review of randomized controlled trials,” *Neuroscience & Biobehavioral Reviews,* vol. 129, pp. 1-16, 2021.

[22] E. Kim, L. Berkovits, E. Bernier, D. Leyzberg, F. Shic, R. Paul and B. Scassellati, “Social robots as embedded reinforcers of social behavior in children with autism,” *Journal of Autism and Developmental Disorders,* vol. 43, pp. 1038-1049, 2013.

[23] J. Nakanishi, H. Sumioka and H. Ishiguro, “A huggable communication medium can provide sustained listening support for special needs students in a classroom,” *Computers in Human Behavior,* vol. 93, no. April, pp. 106-113, 2019.

[24] D. Ricks and M. Colton, “Trends and considerations in robot-assisted autism therapy. Robotics and Automation,” *2010 IEEE International Conference on Robotics and Automation,* pp. 4354-4359, 2010.

[25] H. L. Bradwell and e. al., “Design recommendations for socially assistive robots for health and social care based on a large scale analysis of stakeholder positions: Social robot design recommendations,” *Health Policy and Technology,* vol. 10, no. 3, 2021.

[26] E. Ackerman, “Ollie the Baby Otter Is a Therapy Robot That's Actually Affordable,” 25 March 2015. [Online]. Available: https://spectrum.ieee.org/mit-ollie-the-baby-otter-therapy-robot. [Accessed 26 March 2022].

[27] M. Mori, “Bukimi no tani [the uncanny valley],” *Energy,* vol. 7, pp. 33-35, 1970.

[28] MIT, “Huggable: A social robot for pediatric care,” [Online]. Available: https://www.media.mit.edu/projects/huggable-a-social-robot-for-pediatric-care/overview/. [Accessed 15 March 2022].

[29] B. NT, “Top 5 Best Pet Therapy Robots You Should Know In 2022,” 10 January 2021. [Online]. Available: https://roboticsbiz.com/top-5-pet-therapy-robots-available-in-the-market/. [Accessed 18 October 2021].

[30] M. Cooney, S. Nishio and H. Ishiguro, “Affectionate Interaction with a Small Humanoid Robot Capable of Recognizing Social Touch Behavior,” *ACM Transactions on*

Interactive Intelligent Systems, vol. 4, no. 4, pp. 1-32, 2015.

[31] N. Kleawsirikul and H. M. a. S. Hasegawa, "Unsupervised embrace pose recognition using K-means clustering," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN),* Lisbon, Portugal, 2017.

[32] V. U. Brussel, "Probo - A huggable robotic friend," 2009. [Online].

[33] J. Lee, H. Takehashi, C. Nagai, G. Obinata and D. Stefanov, "Which Robot Features Can Stimulate Better Responses from Children with Autism in Robot-Assisted Therapy?," *International Journal of Advanced Robotic Systems,* vol. 9, no. 3, 2012.

[34] C. N. Y. Y. H. Kozima, "Children-robot interaction: a pilot study in autism therapy," *Progress in Brain Research,* vol. 164, pp. 385-400, 2007.

[35] R. Simut, C. Pop, B. Vanderborght, J. Saldien, A. Rusu, S. Pintea, J. Vanderfaeillie, D. Lefeber and D. David, "The huggable social robot probo for social story telling for robot assisted therapy with ASD children," *Proceedings of the 3rd International Conference on Social Robotics,* pp. 97-100, 2011.

[36] J. M.Beer, K. R.Liles, X. Wu and S. Pakala, "Chapter 15 - Affective Human–Robot Interaction," *Emotions and Affect in Human Factors and Human-Computer Interaction,* pp. 359-381, 2017.

[37] S. Yohanan and K. MacLean, "The Role of Affective Touch in Human-Robot Interaction: Human Intent and Expectations in Touching the Haptic Creature," *Int J of Soc Robotics,* pp. 163-180, 2012.

[38] K. Altun and K. E. MacLean, "Recognizing affect in human touch of a robot," *Pattern Recognition Letters,* vol. 66, pp. 31-40, 2015.

[39] S. a. S. K. D. a. G. S. a. O. B. a. A. L. a. S. N. a. F. K. a. G. H. a. L. D. a. W. P. a. B. C. Jeong, "Designing a Socially Assistive Robot for Pediatric Care," *Proceedings of the 14th International Conference on Interaction Design and Children,* p. 387–390, 2015.

[40] L. Capineri, "Piezoresistive Sensors Fabricated with Conductive Textiles for Monitoring the Step Rate with Read-Out Electronics and Wireless Connection to a Smart Watch,"

Fashion Technol Textile, vol. 3, no. 3, 2015.

[41] F. Yan, A. M. Iliyasu and K. Hirota, “Emotion space modelling for social robots,” *Engineering Applications of Artificial Intelligence,* vol. 100, 2021.

[42] M. H. Mahoor, “AffectNet,” [Online]. Available: http://mohammadmahoor.com/affectnet/. [Accessed 15 3 2022].

[43] R. Verma, “fer2013,” [Online]. Available: https://www.kaggle.com/datasets/deadskull7/fer2013. [Accessed 15 3 2022].

[44] E. a. Z. C. a. C. F. C. a. Z. Z. Barsoum, “Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution,” *ACM International Conference on Multimodal Interaction (ICMI),* 2016.

[45] IEEE, “Furby,” 2021. [Online]. Available: https://robots.ieee.org/robots/furby/. [Accessed 18 October 2021].

[46] A. BILOGUR, “Full batch, mini-batch, and online learning,” 2018. [Online]. Available: https://www.kaggle.com/code/residentmario/full-batch-mini-batch-and-online-learning/notebook. [Accessed 5 March 2022].

[47] J. Brownlee, “How to use Learning Curves to Diagnose Machine Learning Model Performance,” 27 February 2019 . [Online]. Available: https://machinelearningmastery.com/learning-curves-for-diagnosing-machine-learning-model-performance/. [Accessed 24 March 2022].

[48] P. Burgess, “Animated Snake Eyes Bonnet for Raspberry Pi,” 11 January 2017. [Online]. Available: https://learn.adafruit.com/animated-snake-eyes-bonnet-for-raspberry-pi. [Accessed 15 March 2022].

[49] “Mixkit,” [Online]. Available: https://mixkit.co/. [Accessed 18 March 2022].

[50] J. Sardin, “BigSoundBank,” 2005. [Online]. Available: https://bigsoundbank.com/. [Accessed 18 March 2022].

[51] “SketchBook,” [Online]. Available: https://www.sketchbook.com/.

[52] S. Joseph, “Cat Purr,” 2005. [Online]. Available: https://bigsoundbank.com/detail-0436-cat-purr.html. [Accessed 18 March 2022].

[53] Mixkit, “Dog whimper sad,” 2022. [Online]. Available: https://mixkit.co/free-sound-effects/dog/. [Accessed 18 March 2022].

[54] H. Mase, Y. Yoshida and T. Yonezawa, “An interactive stuffed-toy device for communicative description on Twitter,” *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS)*, pp. 1360-1363, 2014.

Appendix

‘dataProcessing.m’

```
%% Load in FER-2013 database
data = readtable('fer2013.csv');

%% Reorangise data
[rows, columns] = size(data);
emotion_cat = ones(rows, 1);
first_pixel = emotion_cat;

for i = 1:rows
 emotion_cat(i) = str2double(data.Var1{i}(1));
 first_pixel(i) = str2double(data.Var1{i}(3:end));
end

VarNames = {'Emotional cat', 'Zeros', 'First pixel'};
T1 = table(emotion_cat, zeros(rows, 1), first_pixel, 'VariableNames', VarNames );
l = data(:, 2:end);

datareorg = [T1, l];

%% Sorting rows in to square matrices.
% data processing has been broken down into chunkcs to make it more
% manageable given memory limitations.

%initiate array
images5000 = cell(1, 5000);

%%

for i = 1:ceil(rows/20)
    a = datareorg{i, 2:end-1}; % collect relevant row
    b = reshape(a, 48, 48)'; % resize into square matrix
    images5000{i} = [b(:, 2:end) b(:,1)]; % assign to cell array
    i % counter variable to display progress
end


%%
for i = floor((rows/20)+1):ceil(2*rows/20)
    a = datareorg{i, 2:end-1};
    b = reshape(a, 48, 48)';
    images5000{i} = [b(:, 2:end) b(:,1)];
    i
end

%%
for i = floor((2*rows/20+1)+1):5000
    a = datareorg{i, 2:end-1};
    b = reshape(a, 48, 48)';
    images5000{i} = [b(:, 2:end) b(:,1)];
    i
```

```
end

%and so on in fraction of rows/20, initiating a new array every 5000 values.


%% Collate all arrays into singular array

imagestotal = [images5000 images10500 images15500 images21000 images26500 images31500
imagesend];


%% Find the indices identifying which images belong to which categories
index0 = [find(emotion_cat == 0)]';
index1 = [find(emotion_cat == 1)]';
index2 = [find(emotion_cat == 2)]';
index3 = [find(emotion_cat == 3)]';
index4 = [find(emotion_cat == 4)]';
index5 = [find(emotion_cat == 5)]';
index6 = [find(emotion_cat == 6)]';
%% Assign images to their relevant emotional categories

emotion0 = imagestotal(index0);
emotion1 = imagestotal(index1);
emotion2 = imagestotal(index2);
emotion3 = imagestotal(index3);
emotion4 = imagestotal(index4);
emotion5 = imagestotal(index5);
emotion6 = imagestotal(index6);

%% Save images as png files to their respective folders

folderName = '.\images\Emotion4';
fig = figure('visible', 'off'); %stopping new figures populating screen
map = [0 255]; %set up pixel value map to ensure imshow works as expected

%% Emotion 0
folderName = '.\images\Emotion0'; %file to be saved to

for i = 0:length(emotion0)
  imgName = sprintf("emotion0_%d", i); %setting up file name with counter
  img = imshow(emotion0{i}, map); % turning matrix into visible image
  saveas(img, fullfile(folderName, imgName), 'png') % saving as a png file to relevant folder
end

%% Emotion 1
folderName = '.\images\Emotion1';

for i = 0:length(emotion1)
  imgName = sprintf("emotion1_%d", i);
  img = imshow(emotion1{i}, map);
  saveas(img, fullfile(folderName, imgName), 'png')
end

%% Emotion 2
folderName = '.\images\Emotion3';
```

```
for i = 0:length(emotion2)
   imgName = sprintf("emotion2_%d", i);
   img = imshow(emotion2{i}, map);
   saveas(img, fullfile(folderName, imgName), 'png')
end

%% Emotion 3
folderName = '.\images\Emotion3';

for i = 0:length(emotion3)
   imgName = sprintf("emotion3_%d", i);
   img = imshow(emotion3{i}, map);
   saveas(img, fullfile(folderName, imgName), 'png')
end

%% Emotion 4
folderName = '.\images\Emotion4';

for i = 0:length(emotion4)
   imgName = sprintf("emotion4_%d", i);
   img = imshow(emotion4{i}, map);
   saveas(img, fullfile(folderName, imgName), 'png')
end

%% Emotion 5
folderName = '.\images\Emotion5';

for i = 0:length(emotion5)
   imgName = sprintf("emotion5_%d", i);
   img = imshow(emotion5{i}, map);
   saveas(img, fullfile(folderName, imgName), 'png')
end

%% Emotion 6
folderName = '.\images\Emotion6';

for i = 0:length(emotion6)
   imgName = sprintf("emotion6_%d", i);
   img = imshow(emotion6{i}, map);
   saveas(img, fullfile(folderName, imgName), 'png')
end

%% Loading images into image data store

ids = imageDatastore(% file address of images \Emotion*.png','LabelSource','foldernames');
[idstrain,idsvalid] = splitEachLabel(ids,0.7); %seperation of data into training and validation sets
auidstrain = augmentedImageDatastore([48 48], idstrain, 'ColorPreprocessing', 'gray2rgb'); %preprocessing of data
auidsvalid = augmentedImageDatastore([48 48], idsvalid, 'ColorPreprocessing', 'gray2rgb');
```

‘liveActionResponse.m’

```
%% Visual and Audio Script

%% Load Files

%load in visual files
restingopen = imread('restingopen.png');
happyopen = imread('happyopen.png');
happyhalfblink = imread('happyhalfblink.png');
happyblink = imread('happyblink.png');
halfsad = imread('halfsad.png');
sad = imread('sad.png');
sadblink = imread('sadblink.png');
happyhalf = imread('happyhalf.png');
fullyhappy = imread('fullyhappy.png');

%load in audio files
[whimper, Fs] = audioread('whimpering.wav');
samples = [1,3*Fs];
[purr, Fs] = audioread('purring.mp3', samples);

%load in network
netstruct = load('finalnet.mat');
net = netstruct.finalnet;


%% Sequence Demo

figure;
resting(restingopen, happyhalfblink, happyblink)
smile(restingopen, happyhalf, fullyhappy, purr)
pause(3)
smiletorest(happyhalf, restingopen)
resting(restingopen, happyhalfblink, happyblink)
sadface(restingopen, halfsad, sad, whimper)
sadfaceblink(sad, sadblink)
sadtorest(halfsad, restingopen)

%% Live action response

cam = webcam(1);
figure;
resting(restingopen, happyhalfblink, happyblink)
inputSize = net.Layers(1).InputSize;

for i = 1:50
    im = snapshot(cam);
    subplot(1, 2, 1), imshow(im)
    I = imresize(im,inputSize(1:2));
    [label,scores] = classify(net,I);
    title(label)
    if label == 'Happy'
        smile(restingopen, happyhalf, fullyhappy, purr)
        smiletorest(happyhalf, restingopen)
```

```
      else
          sadface(restingopen, halfsad, sad, whimper)
          sadfaceblink(sad, sadblink)
          sadtorest(halfsad, restingopen)
      end
    resting(restingopen, happyhalfblink, happyblink);
    pause(5)
 end


%% Sequence Functions

function smile(x, y, z, noise)
    subplot(1, 2, 2),imshow(x);
    subplot(1, 2, 2),imshow(y);
    pause(0.5)
    subplot(1, 2, 2),imshow(z);
    sound(noise, 36000)

end

function smiletorest(x, y)
        subplot(1, 2, 2),imshow(x);
        subplot(1, 2, 2),pause(0.5)
        subplot(1, 2, 2),imshow(y);
end

function resting(x, y, z)
    subplot(1, 2, 2), imshow(y);
    subplot(1, 2, 2), imshow(z);
    pause(1);
    subplot(1, 2, 2), imshow(y);
    subplot(1, 2, 2), imshow(x);
    pause(2)
end

function sadface(q, r, s, noise)
    subplot(1, 2, 2),imshow(q);
    subplot(1, 2, 2),imshow(r);
    subplot(1, 2, 2),imshow(s);
    sound(noise, 27000)
end

function sadfaceblink(s, t)
    subplot(1, 2, 2),imshow(s);
    subplot(1, 2, 2),imshow(t);
    pause(1);
    subplot(1, 2, 2),imshow(s);
    pause(4);
end

function sadtorest(r, s)
    subplot(1, 2, 2),imshow(r);
    subplot(1, 2, 2),imshow(s);
end
```