KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, KUMASI, GHANA

Personalized Affective Robotic Assistive Technology for Children with Autism

By

Rose-Mary Owusuaa Mensah (BSc. MSc Computer Science)

A Thesis submitted to the Department of Computer Science, Faculty of Computational and Applied Sciences, College of Science in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

SEPTEMBER, 2020

DECLARATION

I hereby declare that this submission is my own work towards the PhD and that, to the best of my knowledge, it contains no material previously published by another person, nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in text.

Rose-Mary Owusuaa Mensah …………………… …………………………
PG6852816 Signature Date

Certified By
Prof. James Ben Hayfron-Acquah …………………. …………………………
Supervisor Signature Date

Certified By
Prof. Michael Asante …………………. ……………………………
Supervisor Signature Date

Certified By
Prof. James Ben Hayfron-Acquah …………………. ……………………………
Head of Department Signature Date

ABSTRACT

Social robots are gradually becoming an integral part of human livelihood and have achieved significant results in healthcare, education and entertainment. Recently, research has geared towards identifying ways that such robots can be harnessed in special needs education for children with autism spectrum disorder (ASD). Autistic children have deficits in social interaction, communication and often portray repetitive behaviours. Although technology-based intervention strategies for autistic children could promise great results, many autistic children from resource-constrained environments have been left behind due to the cost implications and technical requirements associated with robotassisted learning. This thesis focused on investigating the suitability of a humanoid robot as an assistive technology for Ghanaian children with autism and proposing strategies for personalization of robot-mediated learning sessions. An iterative prototyping approach was used to design and develop a novel low cost humanoid robot, RoCA, which has been used together with another robot Rosye in a series of empirical studies to assess the initial reaction of some autistic children towards the robots and the engagement levels of the children through longitudinal studies. Results from the preliminary and longitudinal studies indicated that the robot was able to engage some of the children in imitation and general tasks and also succeeded in persuading some children to perform the robot's requests via multiple prompted cueing. The thesis also presents a deep fuzzy framework for personalized affective robot assisted learning in autistic childrobot interactions. This framework is based on a proposed deep learning model,

SingleShot Emotion Detector (SED) and a fuzzy based engagement prediction engine which can use scores, IQ levels and task difficulty as input variables for estimating the engagement levels of autistic children while interacting with social robots. The framework was implemented in the humanoid robot RoCA and another empirical study was conducted to evaluate the effects of the personalization approach provided by the deep fuzzy framework on learning gains in autistic child – robot interactions. Statistical significance of improved learning gains associated with the deep fuzzy approach adopted by RoCA was confirmed by Mann Whitney tests. The thesis also investigated the behavioural intention of special needs teachers to use robots in the classroom to teach autistic children using Unified Theory of Acceptance and Use of Technology (UTAUT) as research model. The results indicated that performance expectancy, effort expectancy and social influence positively affect the behavioural intention of special needs teachers to use robots to teach children on the autism spectrum.

ACKNOWLEDGEMENT

To the Almighty God who although as old as He is, still remains the same, I am very grateful for the gift of life, knowledge and grace bestowed on me to pursue this research. I would like to thank my supervisors, Prof. James Ben Hayfron-Acquah and Prof. Michael Asante for their invaluable support, advice and guidance throughout this PhD. My sincerest appreciation also goes to all the lecturers at the Department of Computer Science for their encouragement and suggestions. I am extremely grateful to my husband, Christian Gyening, my children Alvin, Benedict and Benedicta, my parents and my siblings especially Mr. Samuel Mensah, who has recently passed on to eternal glory. Thank you to Yesutor Gbewonyo and Mr. Samuel Asiedu for their advice and guidance. A massive thank you also goes to all the staff and students at Autism Awareness Care and Training Center, HopeSetters Autism Center, Life Community School and Garden City Special School for their participation in making this work a success. May God bless us all.

DEDICATION

I dedicate this work to all Ghanaian children on the autism spectrum and their caregivers. I hope this work inspires and leads to more re search on technological interventions for special needs education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACRI	Autistic child-robot interaction
ASD	Autism Spectrum Disorder
CNN	Convolutional Neural Networks
DL	Deep Learning
EE	Effort Expectancy
FC	Facilitating conditions
FCL	Fully connected layer
FER	Facial emotion recognition
GA	General activities
HRI	Human-robot interaction
JA	Joint attention
ML	Machine Learning
PE	Performance Expectancy
RL	Reinforcement Learning
RoCA	Robot for children with autism
SED	SingleShot emotion detector
SI	Social Influence
SL	Supervised learning
TAM	Technology acceptance model
ToM	Theory of the mind
UN	United Nations
UTAUT	Unified Theory of Acceptance and Use of Technology
WoZ	Wizard of Oz

CHAPTER 1
INTRODUCTION

Robots ranging from simple and easy to program systems such as Lego® Mindstorms™, to highly sophisticated humanoids are being integrated into educational, health, industrial, entertainment and household activities to enhance productivity. These robots typically vary in levels of intelligence and task delegations. Whereas some are used for highly complex procedures, others commonly known as social robots serve as assistive systems for stress relief and social interactions. Among the special group of people to benefit from social robotic systems are children living with autism spectrum disorder.

Autism Spectrum Disorder (ASD), simply known as autism, is a neurological and developmental disorder that comprises multiple disorders and people diagnosed with this condition exhibit deficits in social interactions, motor skills impairments, language development delays and imagination problems (Cabibihan et al., 2013). This condition is termed a spectrum disorder because there is a range of uneven development in the areas of social communication and interaction (Fuentes et al., 2012). Autism is a lifelong condition; currently, 1% of people worldwide have been diagnosed with ASD and the disorder can occur among all races and ethnic groups (CDC, 2014). The causes of ASD are unknown and the condition cannot be cured. However, the way an autistic child is managed in the initial stages affects the behavioural characteristics that would transcend into the child's adulthood. Technological interventions such as robots are being researched upon as assistive technology to aid autistic people in their daily activities. Prior research e.g. Brian and Matarić (2012) and Cabibihan et al. (2013) suggest that autistic children may easily familiarize and interact with robots as compared to humans because robots are more predictable, deterministic and behave in the same way under similar set of conditions.

1.1 Problem Statement

Several children with autism living in resource constrained environments are unable to access early behavioural interventions which could improve their conditions. This is largely due to inadequate technological aids, lack of trained expertise to handle these children and the financial implications of therapy sessions. Although there have been a few reported successes on using robots in autism therapy to support caregivers in their care for autistic children, more research needs to be conducted to ascertain how autistic children from diverse cultural backgrounds would respond to robotic technology considering the multifaceted nature of the disorder.

Several studies have highlighted the diversified impact of cultural variables and individual differences on suitability, acceptance and effectiveness of technology (Baker & Hubona, 2010; Ennis-Cole et al., 2013). In autism management, Pitten (2008), Tincani et al. (2009), Conti et al. (2015) and Cassio (2015) have indicated that culture and sociodemographic variables need to be incorporated in the design of technological interventions targeted at managing the disorder. Designing robots for autistic children requires a multidisciplinary approach involving the children, their caregivers, healthcare professionals and designers. A user-centered approach needs to be utilized to understand end user requirements, produce design solutions and evaluate these solutions in the natural setting. Currently, a few robotic systems for autistic children exist worldwide; more so, most of them are in the prototype stages and are not available to be public (Huijnen et al., 2017). The few on the market are costly and impractical for trials and long term studies in low to middle income countries.

In some resource constrained countries, the knowledge and potentials of some children on the autism spectrum who have higher than average intelligence quotients and special abilities are being left untapped. Consequently, there is the urgent need for researchers to investigate how autistic children from such environments would perceive and react to robots as therapy partners and propose strategies for incorporating these assistive technologies into the classroom and clinical settings to augment the efforts of the limited skilled personnel.

Scassellati et al. (2012) present that the amount of exposure (i.e. number of interaction sessions) significantly influences the effects of a robotic technology on a child with autism. This is because autistic children are sensitive to changes in their environment and their routines. Their initial reaction to a novel robot may differ from the behaviour the child may exhibit once the robot becomes familiar (Scassellati et al., 2012). However, most of the existing research geared towards this area are single-day interaction sessions.

In order to assess whether the children's attitudes and reactions towards social robot would be same, better or worse with time, longitudinal studies need to be undertaken to engage the children in repeated interactions with robots. Furthermore, the growing number of research aimed at investigating the suitability of robots in autism have mainly utilized the Wizard-of-Oz (WOZ) approach where the robot is controlled unbeknownst to the child and performs pre-defined behaviours. This technique is limiting because it involves a lot of time, human effort and tends to be “boring” as a robot would not be able to adjust to real-time conditions in the environment. From the perspective of assistive technology, social robots ought to react appropriately and adapt to human needs and real-life scenarios.

Currently, increasing autonomy in autistic-child robot interactions is an area of research which needs to be explored. In order to ensure naturalistic interactions and maximize learning gains, social robots need to move from the WOZ approach and be equipped with some intelligence to be able to detect affective states and engagement levels in real time autistic child – robot interactions (ACRI) in order to personalize learning sessions. However, autistic children are quite unpredictable hence modelling social robots to have exact knowledge of the interaction environment is unrealistic. Fuzzy logic is known for its ability to handle imprecision and uncertainty (Nazemizadeh et al., 2014) and may be suitable for controlling the behaviour of social robots in autism therapy. To achieve some level of autonomy and adaptability, machine learning and fuzzy computational models which utilize multimodal data such as affective states, personality traits, social context and physiological disposition could be developed to estimate emotional states and engagement levels of autistic children in real-time ACRI. Many existing facial emotion recognition (FER) systems are based on traditional machine learning algorithms which rely on hand crafted features. These traditional models are known to generate a lot of output feature sets (Bins & Draper, 2001). As such, extracting relevant features from each processing stage to the next requires a lot of computational resources, causes high latencies and often results in low accuracies for real time robotic applications. A challenge which still remains is to be able to accurately and speedily estimate autistic children's affective state from low resolution cameras onboard social robots without the need for sophisticated cameras and devices such as Microsoft Kinect.

Recently, convolutional neural networks (CNNs) have been used to achieve high level accuracy and speed of prediction in many computer vision tasks (Kahou et al., 2015). Therefore, it has become relevant to also investigate whether CNNs and fuzzy logic can contribute to the maximization of learning gains in ACRI by accurately and speedily predicting the emotional states and engagement levels of autistic children during interaction sessions in order for social robots to readapt their behaviour accordingly to suit the children's needs.

1.2 Aim

The aim of this thesis is to explore the suitability of a humanoid robot as an assistive technology for Ghanaian children with autism and propose strategies for personalization of robot-mediated learning sessions.

Specifically, this research intends to:

i. adopt a participatory approach to design and develop a humanoid robotic assistive technology for autistic children ii. assess the reactions, engagement levels and learning gains of Ghanaian autistic children over extended interactio n periods with humanoid robots.

iii. propose a deep learning affect recognition model and a fuzzy -based engagement estimation framework to equip social robots with some level of autonomy to adapt their behaviour to suit individual autistic child's needs. iv. assess the effects of robot behaviour personalization on learning outcomes in robot mediated teaching sessions for children with autism

v. investigate the behavioural intention of special needs teachers in autism management to use robots as assistive technologies in the classroom

1.3 Research questions

The research seeks to address the following questions:

i. Would Ghanaian children on the autism spectrum experience the effects of the uncanny valley during an initial encounter with a humanoid robot?

ii. Would an "unfamiliar" robot be able to continually engage Ghanaian autistic children in learning activities over an extended period of time?

iii. Can a convolutional neural network (CNN) trained on the custom datasets for affect detection perform accurately and speedily in real-time ACRI?

iv. Would personalization of robot mediated learning sessions ensure increased learning gains as compared to a wizard-of-oz operated robot?

v. Are special needs teachers who teach autistic children willing to use robots in the classroom?

1.4 Significance of study

The United Nations report (United_Nations, 2019) on the 2019 World Autism Day celebrated on 2nd April with the theme "Assistive Technologies, Active Participation", highlighted the challenges facing the autism community. The report indicated that although technological devices continually evolve, major barriers such as high costs, unavailability, lack of awareness and inadequate training continue to hinder the successful deployment and use of assistive technologies. More so, in developing countries, more than 50% of the people with disabilities in need of assistive technologies are unable to access them (United_Nations, 2019).

The unmet growing need for adaptive technological interventions in special needs education coupled with limited studies on robots for autism therapy in middle-to-low income countries motivated this research. The suitability of social robots to therapy and social development of children with autism from different ethnic and cultural backgrounds is an area which requires thorough investigation. Most of research geared towards this area have been undertaken in the Western world with barely any done in Africa. There is a saying that *"if you know one child with autism, you only know one child with autism".*

This adage really emphasizes the diversity in autism manifestation among affected individuals. Therefore, personalization strategies tailored towards individual traits, learning abilities and affective states are key in managing the disorder. Robots have been identified as potential tools for autism therapy hence equipping these artificial agents with some intelligence could further enhance learning outcomes. This research intends to investigate how a humanoid robot can be incorporated into therapy interventions for autistic children in Ghana, where the skilled personnel and caretakers of these children are limited. When proven to be effective, low cost robots could serve as assistive technology to aid in therapy sessions to augment the efforts of caregivers in their quest to provide special education for these children who have unique abilities which are currently underdeveloped.

1.5 Organization of thesis

The thesis is structured into six chapters. Chapter 1 provides a background to the study, problem statement, aims and significance of the study. Chapter 2 reviews existing literature on design considerations for social robotic systems, robots as assistive technology in autism therapy, machine learning as well as theoretical frameworks. Chapter 3 focuses on the methodology for the co-design of the humanoid robot RoCA, the various child-robot interaction experiments conducted, the training of the convolutional neural network and development of the fuzzy-based engagement estimation framework. Chapter 4 presents the implementation of the novel humanoid robot RoCA, results from the preliminary and longitudinal studies involving the robot and some Ghanaian children with autism and the Singleshot emotion detector. The discussion of the results and implications of the research are presented in chapter 5. Chapter 6 presents the conclusion of the thesis and recommendations for future work.

CHAPTER 2

LITERATURE REVIEW

This chapter elaborates on autism spectrum disorder, reviews existing research on robot mediated interventions in autism management and machine learning strategies for emotion recognition in human-robot interaction. Theoretical frameworks utilized in the research are also presented in this section.

2.1 Autism

Autism Spectrum Disorder (ASD), simply known as autism, is a neurological and developmental condition that comprises multiple disorders and people diagnosed with the condition exhibit deficits in social interactions, language development delays and imagination problems (Cabibihan et al., 2013). ASD is described with the term “spectrum” because the skills, disabilities, symptoms and levels of impairments vary widely among individuals (MCA, 2012). Dr. Leo Kanner of Johns Hopkins Hospital presented the first clinical description of autism (Zimmerman, 2008). A child diagnosed with autism may live with the condition throughout his lifetime. Currently 1% of people in the world have been diagnosed with autism and the disorder can occur among people of all races, cultures and socioeconomic backgrounds. (CDC, 2014).

Autism spectrum disorder is more prevalent in boys(1 in 42) than among girls(1 in 189) (Christensen et al., 2016). The prevalence rate of autism in America, Europe and Asia is between 1% and 2% (CDC C. f., 2016). However, studies on the epidemiology of ASD in Africa and other continents is limited and therefore, there is the need for empirical studies in order to ascertain the magnitude of the problem in these regions (Bakare & Munir, 2011).

Currently, there are no medical tests that can detect autism; the disorder can be diagnosed by pediatricians, psychologists and neurologists or in some cases a team comprising of these professionals (Boehm, 2016). Typically, children develop at varied paces. However, parents, caretakers and clinicians can monitor a child's emotional, social and linguistic behaviours in order to ascertain the progress of a child towards developmental milestones required for his or her age. Developmental milestones of children within the first two years of life can be classified under social communication (social interaction, non-verbal communication, relationships and play) and cognitive behaviour (interest, routines, movements and sensory activities). Autistic children show deficits in meeting these milestones (Young & Jean, 2019).

In a child's first year, the absence of some behaviours such as smiling, eye contacts and gestures are usually red flags for autism (NIH, 2015). The emotions and facial expressions of autistic children may not match what they say; some are also not able to regulate their emotions and when frustration sets in, they may exhibit aggressive behaviours such as banging head, biting one's self or pulling their hair (NIH, 2015). Factors which can contribute to the early diagnosis of ASD in children include symptoms severity, socioeconomic status and timely parental recognition of early signs and red flags for autism (Daniels & Mandell, 2014). With the increasing prevalence of ASD, there is the risk of over and under diagnosis and therefore autism rating scales such as Autism

Spectrum Rating Scales™ (ASRS™) (Sam & Naglieri, 2010) and Diagnostic and Statistical Manual of Mental Disorders, DSM V (APA, 2013) have been designed to assist in effective diagnosis of the condition. Although autism is a lifelong condition, early interventions and therapies can help to improve the skills and abilities of affected people. These interventions can make extraordinary differences in a child's wellbeing and a small percentage can lose their diagnosis over time (Helt, 2008).

2.2 Autism in Ghana

Ruparelia et al. (2016) report that there are children living with autism in Ghana. As with many countries on the African continent, people with autism were often given cold shoulders because the populace believed such conditions are caused by multivariate factors such as witchcraft, curse from God and other lesser gods or punishment as a result of what a person or his or her family members did. Until recently, children born with autism and their parents could be cast out of their society and totally ignored by their families (Anthony, 2011).

Rural Integrated Relief Service Ghana (RIRSGH, 2010) reports that another organization, Autism In Ghana estimates 1 out of 87 children in Ghana under the age of 3 suffers this disorder and as in conformance with other statistics worldwide, the condition is more prevalent in boys than girls. The low level of knowledge and awareness about autism in Ghana contributes to the late diagnosis and pursuance of treatment options for affected children. Decades ago, educational facilities and training centers for children with special needs were virtually non-existent. A few private autism centers have now sprung up over the country. The Government also owns special schools which accommodate children with varying disabilities including autism. However, it was difficult to find the official statistics on the number of autism centers in Ghana. Notwithstanding these developments, lack of skilled personnel, technological aids, inadequate funds and societal attitudes continue to hamper the provision of adequate care to people living with autism in Ghana.

2.3 Imitation and joint attention deficits in autism

Imitation involves the ability of a person to mimic the behaviour or actions of another person. Basically, to imitate someone, you need to first of all observe what the person is doing and thereafter, repeat the person's actions. Individuals with autism have difficulty and exhibit deficits in imitation skills (Biscaldi et al., 2015). The imitation ability of an autistic child is an important goal which has a correlation to the child's language development, joint attention, play and social interaction skills (Lowry, 2016). More so, due to the fact that imitation serves as a learning and social function in infants, when children are taught imitation skills early, the tendency for their social skills to improve is also high (So et al., 2016). There are several ways to teach imitation to children with autism. However, research shows that these children find object imitation easier than imitating a human's facial expression, gestures and sounds (Ingersoll, 2008). Objects such as scribble markers, lollipops, shake bells, toys and computing devices have been used to teach imitation to autistic children.

Joint attention (JA) skill refers to the ability of a person to share focus with another person on an object or event (Jenkins, 2019). This very important skill begins from 6 to 18 months in infants and as a person develops over the years, JA skills are refined (Lisa et al., 2015). Forms of JA include eye gaze, pointing to objects, making gestures and turning head to focus on an object. To exhibit joint attention, a child can initiate or respond to a JA task.

For example: A child and parent are playing together, and the parent points and says “hello child, look at the sky. The child is able to look up to the sky and say “hello Mum I see the sky”. In this scenario, the parent initiated the joint attention task and the child responded to the task. The development of joint attention impacts significantly on language acquisition and socio-cognitive development of children (Akechi et al., 2011). The joint attention skill in an infant could differ from that of a toddler; more so, even children of the same age could exhibit differences in their joint attention skills (Mundy & Gomes, 1998). The rate of development of JA skills could be influenced by factors including frequency of interaction between children and their caregivers, impairments like deafness and developmental disorders such as autism (Gernsbacher et al., 2008).

2.4 Social robots

Among the promising interventions for managing autism are social robots. Social robots are usually (semi) autonomous mobile machines equipped with the capability to follow, interact and assist humans in daily activities (KPMG, 2016). Social robots can engage, help, interact with humans and other robots and respond to cues from their environment. They can be used as persuasive machines to influence the emotions and behaviour of humans. These robots have been adopted in various sectors such as entertainment, healthcare and education. A study by

Forlizzi and Bartneck (2004), proposed that a design centered approach needs to be adopted in the development of social robots capable of interacting with humans and their environment. In the field of social robotics, important foci of research include embodiment, personality, empathy, adaption and communication techniques (Hutson et al., 2011). The visual appeal of a robot plays a major role in the ability of the child to sustain interest in continuous interaction with the robot.

Majority of the existing robots have physical shapes which can be categorized as: anthropomorphic, non-anthropomorphic or non-biomimetic. Anthropomorphic robots have humanlike shape and tend to exhibit some human characteristics. Developers of anthropomorphic robots do not necessarily aim at building artificial humans; rather, due to the robots' close resemblance to humans, social interactions can easily be facilitated because humans like to interact the same way with machines as with other humans (Duffy, 2003). However, when creating humanoid robots, designers ought to understand the effects of the uncanny valley phenomenon.

The phrase "uncanny valley" was first proposed by Mori (1970). Uncanny valley is a term in human-robot interaction which suggests that robots which have close resemblance to humans are perceived as disgusting, unnerving, revulsive and creepy. Basically, any object which has near resemblance and exhibits human-like features has the tendency to elicit the uncanny valley effect in humans. Objects such as robots, humanlike dolls for children and 3D computer animations could create an uncanny valley effect. However, the level to which people experience this eeriness varies among individuals based on factors such as pre-exposure to the object or familiarity with similar objects. This effect can be created or reduced by incorporating variations in appearance, movement, voice and sounds.

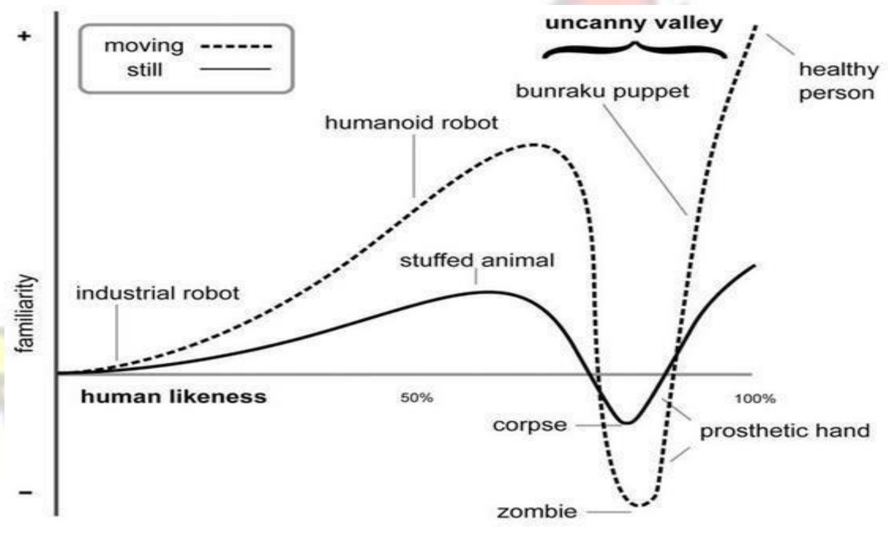


Figure 2.1 Mori's illustration of the uncanny valley effect (MacDorman, et al., 2005)

Figure 2.1 illustrates the process of occurrence of the uncanny valley effect. As robots move very close to human like ness, they tend to be repulsive and undesired behaviours may be exhibited by people towards robots. The uncanny valley phenomenon is an important principle which needs to be considered when developing robots for human use because it could immensely impact acceptance, user experience and interaction.

The second category of robots according to classification by their shapes is nonanthropomorphic robots. These robots are designed to look more like animals and do not resemble humans. Non-anthropomorphic robots are widely used as field and service robots. The third categories of robots, non-biomimetic robots, are those whose shapes do not resemble humans, animals or any biological creatures. The behaviour and functionalities of such robots are not firmly rooted in biological principles. Figure 2.2 show Nao, Keepon and Peekee robots which are examples of anthropomorphic, nonanthropomorphic and non-biomimetic robots respectively.

Nao	Keepon	Peekee
Anthropomorphic robot (NaoRobotics, 2012)	Non-anthropomorphic robot (Kozima & Nakagawa, 2006)	Non-biomimetic robot (Wany_Robotics, 2008)

Figure 2.2 Anthropomorphic, non-anthropomorphic and biomimetic robots

2.5 Robots in Autism Therapy

Robots can be used for repetitive tasks and can be programmed to behave in the same manner under the same set of conditions, making them likely good companions for autistic children who do not respond well to sudden changes in their environment (Cabibihan et al., 2013; Brian et al., 2012; Robins et al., 2004). A robot functions in a limited number of ways and is quite predictable as compared to humans, hence the probability of the child being drawn to a robot is quite high. Robots in autism therapy can take on roles not limited to the following: therapeutic tool, playmate, social mediator, and model social agent (Dautenhan, 2003).

Past research by Dautenhahn & Billard (2002) and Robins (2005) indicate that socially assistive robots (SARs) promise to impact immensely in the therapeutic process, because autistic children easily familiarize and interact with robot companions than humans. However, Scassellati et al. (2012) presents that the amount of exposure (i.e. number of interaction sessions) can significantly influence the effects of a robotic technology on a child with autism. This is because autistic children are sensitive to changes in their environment and their routines. Their initial reaction to a novel robot may differ from the behaviour the child may exhibit once the robot becomes familiar (Scassellati et al., 2012).

Experiments presented in literature on robot mediated therapy for autism management are either single interaction sessions or longitudinal repeated exposures. Single session experiments enable researchers to determine the initial reaction of autistic children to social robots and the effects of robot designs on acceptance by these children. Examples of autistic child – robot single session interactions research are Michaud and Caron (2002), Stanton et al (2008), Werry et al. (2001), Kim et al. (2013), Robins et al. (2009), Valadão et al. (2016), Shamsuddin et al. (2012) and Scassellati (2005). Michaud and Caron (2002) conducted experiments involving a mobile robot, Roball and autistic children with results indicating that the robot caught the attention of the children. The outcome of a study by

Stanton et al. (2008) where autistic children interacted with a robotic dog AIBO suggests that the children spoke a lot more to AIBO as compared to a simple mechanical dog. Werry et al. (2001) report from a study that autistic children who interacted with both a robot and a non-animated toy had more physical contacts, interactions and eye gazes with the robot as compared to the toy. Kim et al. (2013) compared the effects of interactions of autistic children with a social dinosaur robot, PLEO, against interactions with a human or a touchscreen computer game. They observed that participants who interacted with the robot were more verbal during the interaction sessions as compared to the children paired with humans and the computer.

Robins et al. (2009) report from an experiment that a social robot KASPAR was able to serve as a mediator between autistic children and adults present during interaction sessions. Another research by Valadão et al. (2016) indicate that the robot MARIA was able to elicit social skills in some autistic children and these children had more physical contact with the robot compared to the control group. More so, findings from a pilot study by Shamsuddin et al. (2012) in a single session child-robot interaction indicated that four out of the five autistic children used in the study portrayed less autistic behaviour with respect to communication when interacting with the robot Nao. An experiment conducted by Scassellati (2005) with the ESRA robot programmed to perform a roughly 2 minutes short “script” indicated that the autistic children involved were happy and tolerated the robot really well as compared to a human instructor although the robot had no sensory interaction and learning capabilities.

Unlike single session experiments, repeated exposure of the children to a robot is likely to increase familiarity and reduce the influence of variables such as the "novelty effect" (Robins et al., 2004).Such studies are however susceptible to factors such as mood swings and conditions in the environment where the experiments take place. Longitudinal studies in autistic child-robot interactions have been performed over several weeks with typically a maximum of five interaction sessions per child (Scassellati et al., 2012). Some longitudinal studies of autistic child(ren)-robot interaction are Robins et al. (2004), Kozima et al. (2007), Duquette et al. (2008), Wainer et al. (2010) and Valentina (2017).

Robins et al. (2004) discovered from repeated interactions that a social robot can be a salient mediator of joint attention in children with autism. Kozima et al. (2007) conducted interaction sessions among Keepon robot and autistic children for some years and realized that robots are likely to facilitate social interactions among children. Duquette et al. (2008) present that during multiple autistic child-robot interactions, the children paired with the robot Tito exhibited increase in shared focus attention as compared to those paired with a human instructor. Wainer et al. (2010) report from a longitudinal study that, autistic children played more and regarded the robot KASPAR as a partner as compared to their play sessions with a human. A study by Valentina (2017) involving an autistic child and a social robot, IROMEC reported an increase in child-robot eye contact and interaction as compared to that of the interaction between the child and his teacher. The Aurora project also studied how mobile robots can serve as therapeutic tools to aid children on the spectrum (Dautenhahn & Iain, 2004).

Autistic children have joint-attention deficits and therefore some early proponents of robot assisted therapies for autism were geared towards improving joint-attention in autistic children (Robins et al., 2004). In the robot Keepon project, operators manipulated the robot to point to the direction of a child's gaze or towards an object with the "hope" of catching the child's attention (Kozima & Nakagawa, 2006). Other studies that aimed at improving joint attention of autistic children using robots are Robins and Dubowski (2006) and De Silva et al. (2009). Wainer at al. (2014) evaluated how a humanoid robot, KASPAR, can be used to facilitate the playing of games between pairs of children with autism. Detailed observations from their research showed that the autistic children exhibited improved social behaviours after a robot participated in a triadic game session among the pairs of autistic children.

Although these single day and longitudinal studies indicate that robots could be valuable aids in autism therapy, little has been done in the area of harnessing social robots in the classroom to aid academic growth of autistic children. Huijnen et al. (2017) also confirm that social robots have still not been widely assessed in autism education in classrooms. Research needs to focus on how such robots can be used to teach social, academic imitation and joint attention skills to the children.

More so, Thill et al. (2012) indicate that current research in robot assisted autism therapy needs to delve into how the workload on therapists can be reduced by equipping social robots with supervised autonomy. There is the need for high level robotic platform independent models which are capable of inferring the internal states of the children to be able to react appropriately (Thill et al. 2012). Due to the emergence of modern technological advances such as deep learning, it is timely to research into how social robots can sense data from their interaction environment, model the data and readapt their behaviour accordingly to create a friendly learning environment in order to sustain the interest of the children.

2.6 Impact of culture on acceptance of technological interventions for autism therapy

People speculate that autism could be caused in children who come from highly intelligent families, by environmental factors and religious beliefs (Okeke, 2016) but many of these claims have not been verified. Research by Dyches at al. (2004) indicates that culture could play a major role in the diagnosis and treatment options for people on the autism spectrum. It is well known that autistic behaviour varies among individuals but there is currently limited knowledge on multi-cultural variations in autism manifestation (Dyches et al., 2004).

Behaviours of children which may be considered red flags for autism in one country may be completely normal in other environments. Culture could also influence how autistic children interact with their families and outsiders. Similarly, the acceptance of treatment plans for people with autism by their family and autism care centers could differ based on religious beliefs about the etiology of autism, ethnic backgrounds and other cultural variables (Ennis-Cole et al., 2013). Therefore, it is important to factor in culture when designing technological interventions for autism therapy (Pitten, 2008; Cascio, 2015; Tincani et al., 2009). A study by Libin and Libin (2004) where subjects from America and

Japan interacted with a robotic cat indicated that the participants reacted to the robot differently and these variations could include factors such as cultural background. Rudovic et al. (2017) also identified variations in levels of engagement and task execution times among two groups of people from different cultures, specifically Asia(Japan) and Europe(Serbia) during their interactions with the robot Nao.

Currently, most of the studies on robots in autism therapy have largely been focused on Europe and America. The autistic children who have partaken in robot mediated learning experiments are mostly from developed countries (Samadi & McConkey, 2011) such as UK (Wainer et al., 2010; Robins et al., 2004), USA (Stanton et al., 2008; Feil-seifer & Viterbi, 2009), Germany (Robins et al., 2009), Canada (Duquette et al., 2008) and Japan (Lee et al., 2012). A few have been undertaken in developing countries such as Brazil (Valadão, et al., 2016).

Blacher and Mink (2004) point out that, cultural sensitivity should be considered when importing knowledge and cultural practices from one culture to another. Culture could also influence how people react to, accept and interact with technology. Interventions developed and experimented with autistic children in the Western world need to be tested with participants from diverse cultures and resource constrained environments rather than presuming that these technologies would be adequate for children with special needs from these areas. Consequently, there is the need for indigenous research on robot assisted therapy for children from developing countries in order to obtain information about cultural similarities, variations and parameters which could potentially account for acceptance or rejection of robots as partners in autism therapy. Research also needs to throw more light on how pre-exposure to technological gadgets and other robots could influence the acceptance, engagement and learning outcomes of the use of robots in autism therapy.

2.7 Ethical concerns

In human-robot interactions, ethical concerns are a paramount issue and therefore, Syamimi et al. (2014) recommends layers of research protocol which should be considered before commencing robot intervention programs in order to make adequate preparations and obtain consent from all the stakeholders involved. An ethical issue in robots for autism therapy lies in the degree of autonomy given to some robots (Coeckelbergh et al., 2015). The more autonomous a robot is, the less control humans have over its behaviour.

Therefore, who bears responsibility for a robot's action? Can parents trust the robots used in the therapy? Other ethical issues of concern are data privacy and security: will the data obtained from human-robot interaction be stored, how will the data be used and who will use the data?

According to Coeckelbergh et al. (2015), stakeholders approve of using robots in therapy for children with ASD but prefer the activities of these robots are controlled to some extent by therapists. More so, parents, caretakers, therapists and all stakeholders need to be properly informed about the details on how information collected during therapy sessions would be handled. Researchers should also seek parental consent for the participation of their children at all stages of the research.

2.8 Machine learning (ML)

Machine learning is a sub field of computer science which involves equipping computers with the ability to emulate human intelligence by learning from experience without explicit programming (Naqa & Murphy, 2015). Machine learning techniques have been utilized in diverse fields such as computer vision, pattern recognition, medical sciences, email filtering and entertainment. Common machine learning algorithms include supervised learning, reinforcement learning and unsupervised learning (Ayodele, 2010). Supervised learning (SL) basically implies learning from examples; a supervised algorithm is supplied with a training set and a test set. Each training set consists of *n* ordered pairs (x_1,y_1), (x_2,y_2), … (x_n,y_n) where each x_i is a measurement and y_i denotes the label for that particular data point. The test set contains a set of *m* measurements without output labels: $(x_{n+1}, x_{n+2}, \ldots X_n+_m)$ (Learned-Miller, 2014). A common formulation of SL is in classification problems where the algorithm is required to learn a function mapping a vector into one out of several classes by looking at many input-output examples of the function (Ayodele, 2010). Examples of supervised learning algorithms in classification are neural networks, perceptron algorithms, support vector machines, decision trees, random forests and logistic regression (Ayodele, 2010). The type of supervised learning algorithm to use depends on the application domain and the nature of the dataset. For example, logistic regression is best suited for regression problems since it is more robust to noise and can also interpret output as probabilities (Bi & Jeske, 2010). However, logistic regression algorithms are not able to handle categorical features.

Decision trees can be used for both classification and regression tasks and are applicable to domains such as fault detection (Yang et al., 2009) and intrusion detection (Stein et al., 2005). Decision trees are less effective when working with continuous variables. Support Vector Machines (SVMs) are known to work well in non-linear classifications and complicated domains such as speech recognition (Zhang & Gales, 2012), remote homology (Muda et al., 2011) and texture classification (Anantrasirichai et al., 2013). SVMs have issues with noise, large datasets and choice of kernels.

Neural networks are supervised learning algorithms which mimic the functioning of the human brain. Neural networks can learn on their own and do not require explicit programming. These networks are now being explored in natural language processing (Li et al., 2015), computer vision (Gopalaktishnan et al., 2017), financial predictions (de Oliveira et al., 2013) and visual recognition (Gu, et al., 2018). Neural networks are able to work well with limited data and often achieve accurate results.

Reinforcement learning (RL) is another type of ML algorithm in which the learner is not explicitly told which actions to take but should be able to discover actions which yield the utmost reward (Sutton & Barto, 2015). RL is based on try and error where the learning algorithm begins with random actions (since there is no initial knowledge) and learns to perform the expected output (Ritschel, 2018). Reinforcement learning is being used in diverse areas such as self-driving cars (Liang et al. 2018), industry automation (Meyes, et al., 2017) and health care (Gottesman, et al., 2018). Reinforcement learning requires a lot of training time and multiple iterations hence can be time consuming and less effective in some domains.

In unsupervised learning, an algorithm is given inputs $x_1, x_2, \ldots x_n$ without accompanying labeled data, target outputs or feedback from the environment (Lison, 2015). The algorithm then attempts to find patterns or grouping in data. These algorithms are best situated for instances where desired outcomes are unknown. Applications of such algorithms include clustering, anomaly detection and latent variable models.

2.9 Deep learning (DL)

Deep learning is the new era of machine learning algorithms developed with aim of moving machine learning closer to artificial intelligence. DL algorithms are inspired by and mimic the behaviour of the human brain, by using multi-layered artificial neural networks to learn representations from data (LeCun et al., 2015). Artificial neural networks are one of the fastest performing and accurate machine learning algorithms. However, the amount of data supplied to a model directly impacts the model accuracy. In instances where datasets for training deep learning models are limited, a new technique known as transfer learning can be applied (Weiss et al., 2016). Transfer learning is a technique in which knowledge gained from a learned environment is applied to a new environment (Weiss et al., 2016). In transfer learning, there is a source domain D_s with matching source tasks T_s. A target domain D_t also exists with target tasks T_t (Weiss et al., 2016).

The objective of transfer learning is to improve the predictive function of the target $f_t(.)$ by using information from the source domain D_s and the source tasks T_s. Typically, $D_s \neq D_t$ and $T_s \neq T_t$ (Weiss et al., 2016). DL algorithms can be developed using supervised or unsupervised approaches. The success of machine learning algorithms also depends on selecting an architecture that best fits the problem domain (Miikkulainen et al., 2017). Some architectures are: convolutional neural networks (CNN), recurrent neural networks (RNN) and deep residual networks (DRN). Convolutional neural networks are known to produce remarkable results for computer vision tasks (Wootaek et al., 2016).

2.10 Building blocks of convolutional neural networks

CNNs are deep learning algorithms whose architecture is inspired by the organization of the animal visual cortex (YamashitaRikiya et al., 2018). CNNs are feed-forward deep learning image classification algorithms which can automatically learn spatial hierarchies of features. They consist of three layers namely: convolutional layers, pooling layers and fully-connected layers as shown in figure 2.3.

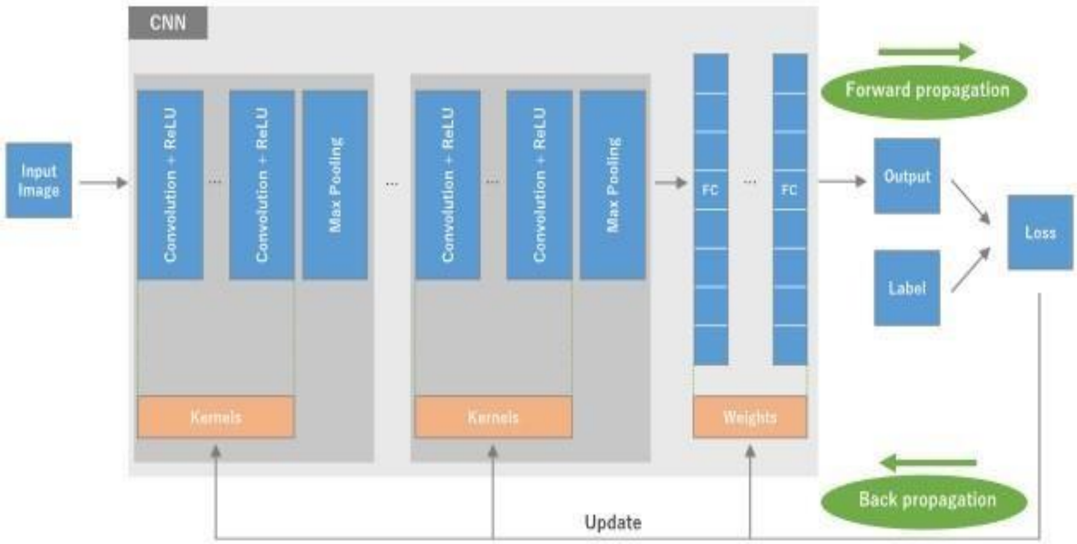


Figure 2.3 An overview of CNNs (YamashitaRikiya et al., 2018)

Convolution layers in CNN extract input features from an image by performing convolution operations. Convolution operations are linear mathematical operations. To execute a convolution, a filter or kernel is applied over an input data to comp ute a dot product called the feature map. The size of the feature map depends on the depth (number of filters used), stride (number of pixels by which the filter is strided over the input data) and zero padding (padding input matrix with zeros in order to control the size of the feature map). Convolutions are very essential operations because they lead to sparse connections and parameter sharing. In figure 2.4, a 4x3 matrix is convolved with a 2x2 filter using a stride of 1. The output is a 3x2 matrix.

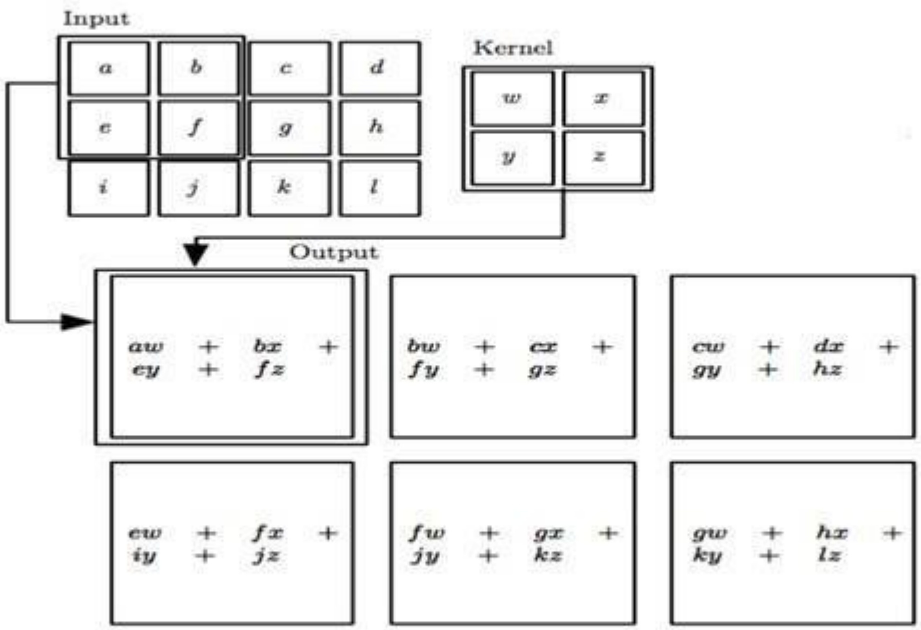


Figure 2.4 Convolution process in deep learning

Rectified Linear Units (ReLUs) are non -linear activation functions performed after convolution operations. ReLUs increase non -linearity in CNN by replacing all negative values in the feature map with zero. Mathematically, the output of a ReLU is defined as $f(x) = \max(0,x)$. Pooling, usually known as spatial spooling or down sampling reduces the dimensionality of the feature map while maintaining important information. Some types of pooling are max pooling and average pooling. Average pooling calculates the average from each selected region of the input matrix and uses the derived averages to create a new matrix. Max pooling, the most preferred method, divides the input matrix into multiple regions and takes the maximum value from each region. In figure 2.5, a 2x2 filter is run over a 4x4 input matrix with stride=2. For each region, the maximum value is picked to create a new matrix of a smaller size.

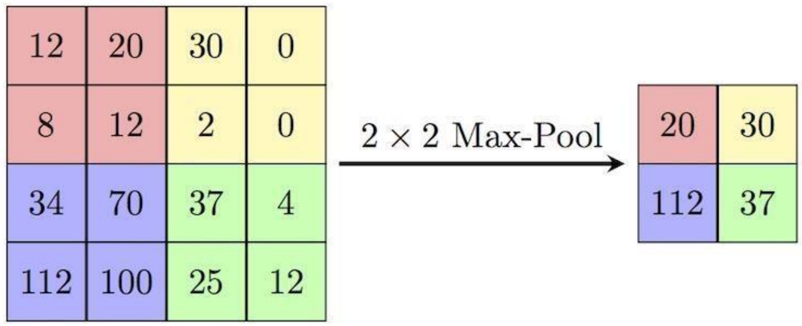


Figure 2.5 Example of max pooling

The output of the final pooling is converted into a 1 dimensional array of high level features and connected to Fully connected Layers (FCL). The FCL is a multilayer perceptron which does high level reasoning in order to predict a classification for the input data. There can be one or more FCL's in a CNN. The FCL takes each neuron from the previous layer and connects it to every neuron in the FCL. The FCL also contains a softmax activation function which assigns probabilities to each class in a multinomial class. In human-robot interaction, robots could take better decisions if they know the kinds of objects in the interaction environment and their exact locations. For this reason, it is timely to explore object detection machine learning models for human-robot interaction.

2.11 Object detection

Object detection is a computer vision technology which involves the process of detecting the presence of objects (e.g. cat, dog, human face, book) in digital images and video sequences, localizing the objects, drawing bounding boxes around them and finally classifying the objects (Zhao et al., 2019). Application areas of object detection include video surveillance, image retrieval, self-driving cars and security systems. Every object has its unique and salient features that characterize and differentiate it from other objects (Fernandez Robles, 2016**)**. Object recognition is an extended form of image classification. Whilst image classification usually classifies an image as belonging to one out of many classes with or without localization, object detection algorithms recognize multiple objects, localize them with bounding boxes and generate the predicted class of each object (Agarwal, 2011**)** , an example shown in figure 2.6.

(a) Image classification: Tomato

(b) Object detection: Apple, Orange, and Pear

Figure 2.6 Differences between image classification and object detection (Agarwal, 2011)

There are two main approaches to object detection using convolutional neural networks, namely the one-stage methods and two-stage methods. In the two-stage method, region proposal networks are used to generate region proposals in the first stage and another network is used to make the final prediction. Models such as Region-Based Fully

Convolutional Neural Networks (R-FCN) (Dai et al., 2016), Fast-Region-based Convolutional Neural Networks (Fast-RCNN) (Girshick, 2015) and Faster-RCNN (Ren et al., 2015) are all two stage methods. The one-stage methods are simple and involve making a fixed number of predictions on a grid. Examples of single-stage object detection algorithms are You-Only-Look-Once (YOLO) (Redmon et al., 2016) and Multibox SingleShot Detector (SSD) (Liu et al., 2016). The one-stage methods do not make use of region proposals and as a result, they are usually faster than the two stage methods (Soviany & Ionescu, 2018). Hence, for some application domains which require fast realtime processing of data, one stage object detection methods may be preferred due to their speed of task execution.

In human-robot interaction, one-stage object detection algorithms can be of more advantage because these algorithms can detect the presence and exact location of objects at a faster rate. In an application area such as robot assisted learning for autistic children, an object detection algorithm can help a robot to know the objects in the interaction environment so as to take appropriate actions if a child is close to a “dangerous” or inappropriate object. Similarly, a fast object detection algorithm can predict the emotional states of autistic children and the robot can utilize that information to decide on the next action to take.

The standard metrics used by image classification algorithms such as True Positive Rate (TPR), False Positive Rate (FPR) and Receiver Operating Characteristic (ROC) curve may not apply to object detection because object detection algorithms may identify many objects of different classes in an input image. More so, the object detection model needs to be able to calculate both localization and classification accuracies and losses. As a result, Average Precision (AP) is the principal metric used to rate the accuracy of many object detection algorithms (Padilla et al., 2020).

Average Precision is calculated based on precision and recall by identifying true positives and false positives (taking into consideration intersection over union (IoU)) as well as false negatives. In object detection, true negatives do not apply because every part of the image which does not contain an object is a negative (Padilla et al., 2020). Precision can be calculated by dividing the total number of true positives by all detections while recall is the total number of true positives divided by all ground truths. To get the average precision of a class, also known as area under the precision-recall curve, precision and recall values for the class have to be plotted on a graph. The mean average precision is calculated by finding the mean of the sum of all the average precisions of the classes (Henderson & Ferrari, 2016).

2.12 Machine learning in robot assisted diagnosis and therapy

Traditionally, autism is diagnosed by health professionals by observing behavioural characteristics of a person. Recently, focus has been placed on research towards diagnosing and providing interventions for autism via machine learning (ML). Much of this effort has been geared towards ML strategies to diagnose autism (Thabtah, 2017).

Crippa et al. (2015) developed a supervised machine learning (SML) algorithm to classify children with low functioning autism based on upper limb movement. Other SML algorithms to predict autism are based on brain regional cortical thickness (Yun, et al., 2010), multi-parametric magnetic resonance imaging data (Zhou et al., 2014) and face processing abnormality (Wenbo et al., 2016).

Notwithstanding, in robot mediated therapy sessions for children with autism, machine learning methods could be used to model and predict the affective states and behaviour of autistic children during interactions. These predictions can further guide social robots to deliver personalized interventions to meet the needs of each child. To achieve the required speed, latency and power efficiency of machine learning implementations for humanrobot interactions, deep learning, a new area of machine learning which utilizes artificial neural networks needs to be explored due to the deficiencies associated with the traditional machine learning approaches.

2.12.1 Affect recognition in human robot interaction (HRI)

Affect is a physiological term used to describe a person's emotion. Automatic detection and classification of human affect ensures "natural" bi-directional communication and enriches human computer interaction (McColl et al., 2016). In special needs education for children with autism who may exhibit deficits in verbal communication, a robot as a therapy partner can be used to sense nonverbal cues from a child and predict his or her affective state in order to readjust the intervention when necessary (Liu et al., 2008).

Facial expressions play a pivotal role in social interactions and can be used to infer a person's emotional state or intention (Levi & Hassner, 2015). Automatic facial emotion recognition (FER) in unconstrained environments such as human-robot interactions still remain an interesting problem in computer vision (Mollahosseini et al., 2015). Automatic detection of emotions is challenging due to computational requirements (Wimmer et al., 2008), technological constraints and real-time processing of detected affective states (Castellano et al., 2010). Data collected by a robot during interactions is also liable to noise such as occlusions and varied illumination (Rouast et al., 2019). More so, just like humans typically do, social robots need to be able to combine multisensory data picked from the environment and interpret them in order to infer affective states as well as engagement levels in order to readapt their behaviour accordingly.

In human-robot interaction, majority of the FER systems proposed for robots have been based on traditional feature extraction and machine learning techniques. Giorgan and Ploeger (2011) extracted Gabor features from faces and used two machine learning algorithms SVM and Adaboost for emotion classification system to be used by domestic robots. Alonso-Martin et al (2013) used OpenCV to detect faces and two third party software systems SHORE and CERT for a FER system running on the robot Maggie. Cid et al. (2013) used Gabor filters as feature extractor and a dynamic Bayesian network as classifier for a robot Muecas to detect people's emotions. Leo et al. (2015) also proposed a FER system based on HOG descriptor and support vector machines (SVM) for deployment on the Robokind™ R25 robot. Lui et al. (2017) developed a FER system based on 2D Gabor filters, LBP and extreme learning classifier for Nao robot.

The traditional algorithms reviewed are based on shallow learning or handcrafted features such as local binary patterns (Shan et al., 2009), Scale Invariant Feature Transform (SIFT) (Lowe, 1999), Histograms of oriented gradients(HOG) (Dalal & Triggs, 2005) and nonnegative matrix factorization (Zhi et al., 2011) which require extensive computations making them unsuitable for real time human-robot interaction applications. More so, traditional machine learning algorithms such as SVMs are known to predict emotions for images captured in controlled lab experiments but have poor generalizability to spontaneous images captured in the wild (Mollahosseini et al., 2015).

For challenging domains where multiple sources of noise are likely to exist, emotion recognition is currently being tackled via deep neural networks (Zhang et al., 2015; Li & Deng, 2018; Cheng et al., 2018). Although deep learning has produced excellent tasks in other application domains, challenges still remain in deep learning for facial emotion recognition. Existing facial emotion recognition systems based on CNN follow a similar pattern as the traditional FER systems by separating the face detection and feature extraction/classification modules. For example, Mayya and Radhika (2016) detected faces via OpenCV and performed feature extraction using a deep CNN on the CK+ and JAFFE datasets for emotion classification. Yu and Zhang (2015) detected faces and classified emotions via ensemble of face detectors and classifiers. This approach of using separate modules for face detection and classification is also not suitable for embedded systems and robots due to the computational resource constraints. Insufficient facial expressions dataset is a major problem which could cause overfitting in deep learning architectures for human-robot interactions.

While object detection has achieved significant speed and accuracy in domains such as airplane detection (Chen et al., 2018), ship detection (Wang et al., 2018), indoor environmental perception for robots (Wang, et al., 2019) and object grasping (Fang &

Zheng, 2019), object detection is yet to be applied to real time affect recognition by social robots in autistic child-robot interactions.

2.12.2 Personalization of autistic child-robot interactions

Every child with autism has his or her unique abilities and weaknesses. Ideally, individualized education plans which incorporate the specific needs of each child are best suited for them. Applying this principle to robot mediated learning in autism, it is imperative for interaction scenarios for social robots to be personalized based on the needs of each child. More so, robots need to be able to sense data such as the emotional states and engagement levels of the children during the interaction sessions to be able to streamline the activity sessions in real-time to maximize learning gains. Not much work has been done on applying deep learning emotion recognition systems to enhance robot mediated therapy for ASD. The few emotion recognition systems targeting this area have been based on traditional machine learning algorithms using datasets of typically developing people and audio (Leo et al., 2015; Kim et al., 2017).

Current challenges which remain are lack of or inadequate domain specific datasets which consequently results in low prediction accuracies of the FER systems for human-robot interaction. To address the problem of limited domain specific dataset of facial expressions, uneven distribution of images, latency issues and computational inefficiency, this thesis proposes a SingleShot emotion detector based on transfer learning on SSDLite, an object detection algorithm to predict the emotional states of an autistic child during ACRI sessions. The proposed model, SingleShot emotion detector is a CNN algorithm for face localization and classification of facial expressions. Unlike traditional emotion recognition systems which separate face detection and classification pipelines making them less suitable for robotic applications due to their high latency, coupling and low frames per second(FPS), the proposed model combines face detection and classification into a single pipeline in order to provide high frames per second(FPS), low latency and efficient power usage for real-time emotion detection tasks. The proposed model would be deployed on a humanoid robot and evaluated in real time to assess the effects of personalization based strategy on learning outcomes.

More so, in order to maximize learning gains, social robots need to be able to detect the engagement levels in real time autistic child – robot interactions (ACRI). Based on a child's current affective state and the estimated engagement level of a child, the robot can then personalize the learning sessions. Engagement estimation in autistic child –robot interaction has not received much attention from researchers (Feng et al., 2017). Current engagement estimation models are typically based on unisensory modalities such as facial images (Rudovic et al., 2018), head pose (Anzalone et al., 2015) and body movements (Colton et al., 2009).

Due to the multifaceted nature of the disorder, the combination of multimodal data such as affective states, IQ levels, learning progress and physiological disposition would better estimate engagement levels of autistic children in real-time. Due to the high level of uncertainty and unpredictability of children on the autism spectrum, this thesis also proposes and evaluates the application of fuzzy logic technology (a technique known to work accurately in uncertain and imprecise environments) using multiple crisp inputs to engagement estimation in autistic – child robot interactions.

2.12.3 Behavioural intention of special needs teachers to adopt robotic assistive technology

The successful integration of robots in educational settings require maximum support from teachers (Mubin et al., 2013). However, much emphasis has been placed on the design of social robots for people with special needs without addressing the perceptions of the teachers about robots and factors which could influence their behavioural intention to adopt robots in educational settings. The behavioural intention of people towards technology could be motivated by the perceived benefits as well as psychological determinants (Wu, 2009).

User acceptance could also be influenced by sociodemographic variables such as culture, age, gender, educational background and previous technological experience (Crabbe et al., 2009; Jain & Rekha, 2017). The general stereotype is that younger people are more likely to accept technology than older people (Flandorfer, 2012) but research by Mitzner et al (2010) contradicts this claim. With respect to gender, Flandorfer (2012) indicates that men appreciate technological devices more as compared to women. A person with higher educational background or previous technological experience is likely to accept new technological device with ease as compared to someone with very little education who is unfamiliar with modern technology (Koç et al., 2016).

Research towards social acceptance of robots can be conducted in the laboratory setting or in the “wild” (in the natural environment). However, due to the fact that these robots are going to assist users on a personal level in everyday life, experiments geared towards understanding variables influencing social acceptance need to be carried out in the natural setting (Weiss et al., 2010). It is difficult to come across studies focusing on how people in developing countries perceive robotic assistive technology. It is therefore important to utilize a multicultural approach to investigate how demographic factors influence the perception and behavioural intentions of teachers to incorporate robots in special needs education.

2.12.4 Technology acceptance modeling frameworks

Two of the well-established theoretical frameworks for modelling the willingness of people to use new technology are Technology Acceptance Model, TAM (Davis, 1989) and Unified Theory of Acceptance and Use of Technology(UTAUT) (Venkatesh et al., 2003). TAM proposes that two main determinants, Perceived Usefulness (PU) and Perceived Ease of Use (PEOU) influence acceptability of new technology. PU is defined by Davis as “the degree to which a person believes that using a particular system would enhance his/her job performance” (Davis, 1989). PEOU is the “degree to which an individual believes that using a particular information technology system would be free of effort” (Davis, 1989). Humans are social factors and as such their environments tend to impact on their beliefs and choices. As such, TAM has been criticized for its simplicity and failure to incorporate the characteristics of individuals and their cultural backgrounds

(Jiang et al., 2010). Extensions to TAM such as TAM2 (Venkatesh & Davis, 2000) and Unified Theory of Acceptance and Use of Technology(UTAUT) (Venkatesh et al., 2003) have been proposed to address the criticisms.

Unlike TAM, UTAUT places emphasis on how social constructs and norms influence people's acceptance of technology. The UTAUT framework (figure 2.7) consists of four principal independent variables: performance expectancy (PE), effort expectancy(EE), social influence(SI), and facilitating conditions(FC) (Venkatesh et al., 2003). Venkatesh et al (2003) identifies four variables: age, gender, experience and voluntariness of use as the moderators of the principal variables (Venkatesh et al., 2003). Behavioural intention is the dependent variable. PE and EE in UTAUT can be mapped to PU and PEOU in TAM respectively (Yogesh et al., 2011). Social influence is defined as the degree to which the beliefs of family, colleagues or peers would influence a person's decision to adopt technology (Venkatesh et al., 2003). Venkatesh et al. (2003) define facilitating conditions as "the degree to which an individual believes that organisational and technical infrastructure exists to support the use of a system".

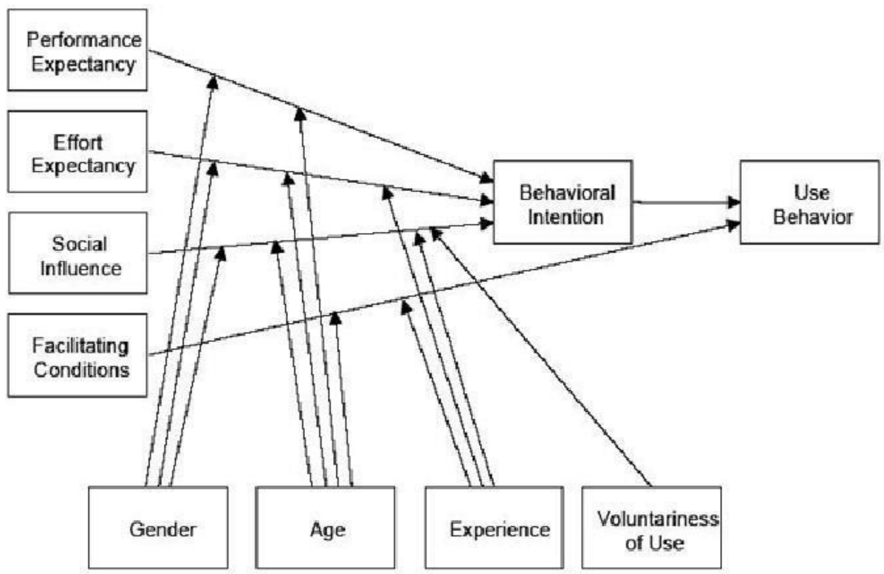


Figure 2.7 The UTAUT Model (Venkatesh et al., 2003)

Conti et al. (2017), Heerink et al. (2006), De Ruyter et al. (2005) have used UTAUT framework to investiga te factors influencing acceptability of robots but none of these studies were conducted with people from developing countries.

2.13 Theoretical Framework

This thesis has been motivated and guided by some theories originating from autism research namely theory of constructivism, scaffolding theory and theory of the mind. The theory of constructivism, a learning theory which deals with the way people learn, acquire and process knowledge could be applied to artificial intelligence. According to Piaget's theory of constructivism, children learn, acquire and reconstruct knowledge by manipulating physical artifacts, observing their behaviour and interacting with their environment (Piaget, 1974). This research seeks to find out whether Ghanaian children on the autism spectrum can learn and acquire new knowledge from the robot through play, observation and interaction sessions.

Scaffolding theory can also be applied in human-robot interaction studies (Robins et al., 2005). Scaffolding refers to manipulating robot behaviour during HRI studies to suit the emotional and cognitive states of the people involved in the interaction. By so doing, the possibility of a child being stressed, agitated and uncomfortable would be minimized. This research also intends to investigate whether manipulating the robot to personalize learning sessions using multivariate parameters can sustain the interest of the children in the robot interaction sessions and maximize learning gains.

According to Pedersen (2018), "Theory of mind (ToM) is the ability to recognize and attribute mental states: thoughts, perceptions, desires, intentions, feelings to oneself and to others and to understand how these mental states might affect behaviour". This theory helps humans to understand why people are likely to act in a certain way and it also helps us to predict how people may act under given circumstances. For example, Ama puts her school bag on the dining table and goes outside to play. Her mother arrives home shortly afterwards, picks Ama's bag and put in a wardrobe in Ama's room. Ama finishes playing and goes to the dining table to look for her school bag so she can take her homework book. Ama falsely thinks her bag is still on the dining hall table because she does not know that her mother moved it. This scenario implies that Ama has theory of the mind based on her knowledge (she knows she kept her bag on the dining hall table) and beliefs.

From the age of 4, children start to develop theory of the mind and therefore they are likely to understand that a person is acting in a certain way due to a false belief and not the reality (Kloo et al., 2010). However, children with developmental delays such as autism may later or never develop theory of the mind (Barron-Cohen, 2000). ToM plays a major role in a child's mental and social development and has also been linked to the ability to understand one's own and other's behaviour, false-beliefs, hold meaningful conversations with people, resolve conflicts and understand the needs of others. Children on the autism spectrum have varying levels of deficits in ToM; they may not know that people have their own plans, beliefs and thoughts and would experience difficulty in understanding the emotions and attitudes of others (Rastall, 2016).

The inability of an autistic child to develop theory of the mind affects their social interaction with humans. ToM can also be linked to how people experience the uncanny valley phenomenon i.e. the tendency to be revulsive, feel creeped out and disgusted by objects which have close resemblance to humans. Therefore, it is crucial to investigate whether an autistic child who may or may not exhibit deficits in ToM would experience the uncanny valley effect during an interaction session with a humanoid assistive technology such a social robot. The results obtained would better inform design choices in terms of appearance for the construction of robots for use by children on the autism spectrum.

2.14 Conclusion

Whereas the benefits of social robots can be enormous, there are many factors hampering the successful implementation of robots in autism therapy. There is limited research on the categories of shapes, sizes and features of robots that would appeal most to children with autism, for e.g. would an autistic child prefer a humanoid to a non-humanoid robot, or vice versa? Would a child prefer to play with a robot smaller than or bigger than him? Would a child be more stimulated by a moving robot or a stationary robot? Another feature which is seemingly missing from most of the robots for autism is the ability to incorporate adaptive inference models which can readapt the robot's behaviour to suit the children's affective states, cognitive abilities and behaviour. Most of the research undertaken in the area of robots for autism therapy has produced more of qualitative data than quantitative making it difficult to statistically assess the progress of the participants over time. Comprehensive review presented in Diehl et al. (2012) indicate that the number of participants in previous robot mediated therapy experiments ranged from 1 to 16.

A lot of the autistic child- robot experiments undertaken were conducted over a short period of time. This could be due to the fact children with autism are classified under "protected groups" and as a result, obtaining consent to engage them in long term studies is difficult. There is the need for large scale long term longitudinal studies involving more autistic children in order to obtain rich quantitative data for analysis. Another issue of concern is the fact that people feel that robots may take over the jobs of humans and therefore, people are reluctant to accept assistive technologies. Perhaps, a more worrying factor is the tendency of autistic children to become emotionally attached to these robots (Coeckelbergh et al., 2016).

Due to the high cost of purchasing or hiring existing robots in autism therapy, autism centers in developing countries may not be able afford them for experiments. More so, assembling, building, programming and controlling social robots are difficult tasks which require high level of skills and expertise. The acceptance, effectiveness and successful use of robots as assistive technologies in ASD therapy depends on many factors such as the design, mobility, capabilities and appeal of the robot (Michaud et al., 2007). Hence it is vital for researchers to adopt a participatory-design approach involving practitioners, teachers and the children to investigate the social norms of the target environment, needs, challenges, capabilities and expectations of the end users in order to design and develop useful tools to aid their activities.

Having identified these gaps in robot mediated therapy for autism management, this thesis focuses on addressing some of the challenges discussed. Specifically, the research seeks to explore the suitability of a humanoid robot as an assistive technology for Ghanaian children with autism and propose strategies for personalization of robot-mediated learning sessions. The objectives involve adopting a participatory design approach involving the caregivers, health care professionals and the children to develop a humanoid robot to aid in the teaching sessions and create appropriate interaction scenarios for robot mediated learning sessions.

As part of the research objectives, a preliminary study would be performed to evaluate the initial reaction of some Ghanaian autistic children towards an unfamiliar robot. Longitudinal studies would be conducted to assess the reactions, engagement levels and learning gains of Ghanaian autistic children over extended interaction periods with humanoid robots. A deep learning model for detecting emotional states of Ghanaian autistic children would be developed and trained on two custom datasets, one containing images of typically developing Ghanaians and the other containing images of autistic children depicting various emotions. A fuzzy-based engagement prediction model would also be proposed to autonomously estimate the engagement levels of the children in real time ACRI. The emotion detector and fuzzy-based models would be incorporated into the developed social robot and tested in real -time autistic child-robot interaction sessions in order to assess the performance of the models and the effects of personalized interactions on learning outcomes in a classroom setting.

CHAPTER 3

METHODOLOGY

In this chapter, an elaboration of the methodology, design and data collection methods utilized in the research are presented. This thesis aimed at exploring the suitability of a humanoid robot as an assistive technology for Ghanaian children with autism and proposing strategies for personalization of robot-mediated learning sessions. The two main methodological approaches adopted were participatory design and experimental research. Most of the research undertaken in the area of robots for autism therapy has produced more of qualitative data than quantitative, making it difficult to statistically assess the progress of the participants over time. Hence an experimental research approach was adopted to collect both quantitative and qualitative data to derive comprehensive knowledge and data which can be used for further research. The main phases of the research involved:

i. Preliminary and longitudinal empirical autistic child-robot interactions using the humanoid robot Rosye ii. Novel robot RoCA's requirements elicitation, interaction scenarios specification and iterative robot development iii. Development of the Singleshot Emotion Detector (SED) based on transfer learning iv. Development of a fuzzy-based engagement prediction model

v. Integrating the SED and the fuzzy models into the developed robot and conducting experiments to assess the effects of personalization on learning gains.

vi. Investigating the behavioural intention of Ghanaian special needs teachers in autism therapy to adopt robotic assistive technology

Two different humanoid robots, Rosye and RoCA were used in the research process.

Rosye is a simplistic multi-coloured adult-size humanoid robot while RoCA is a childlike robot. After the preliminary and longitudinal studies, the caregivers suggested that it would be more appropriate to use a child-sized robot in order for the children to see it as a playmate and also due to colour sensitivity issues, a neutral coloured robot was preferred for future studies. Based on this information and consultation with some stakeholders, a participatory-design approach was adopted to elicit requirements of a new robot from stakeholders, produce the design solutions and evaluate the proposed designs in the classroom setting.

The developed social robot for children with autism has been named RoCA (**Ro**bot for **C**hildren with **A**utism). RoCA was programmed in C# using Visual Studio 2017 IDE and Ez-Script using Ez-Builder. Figure 3.1 indicates the methodological structure of the entire research process.

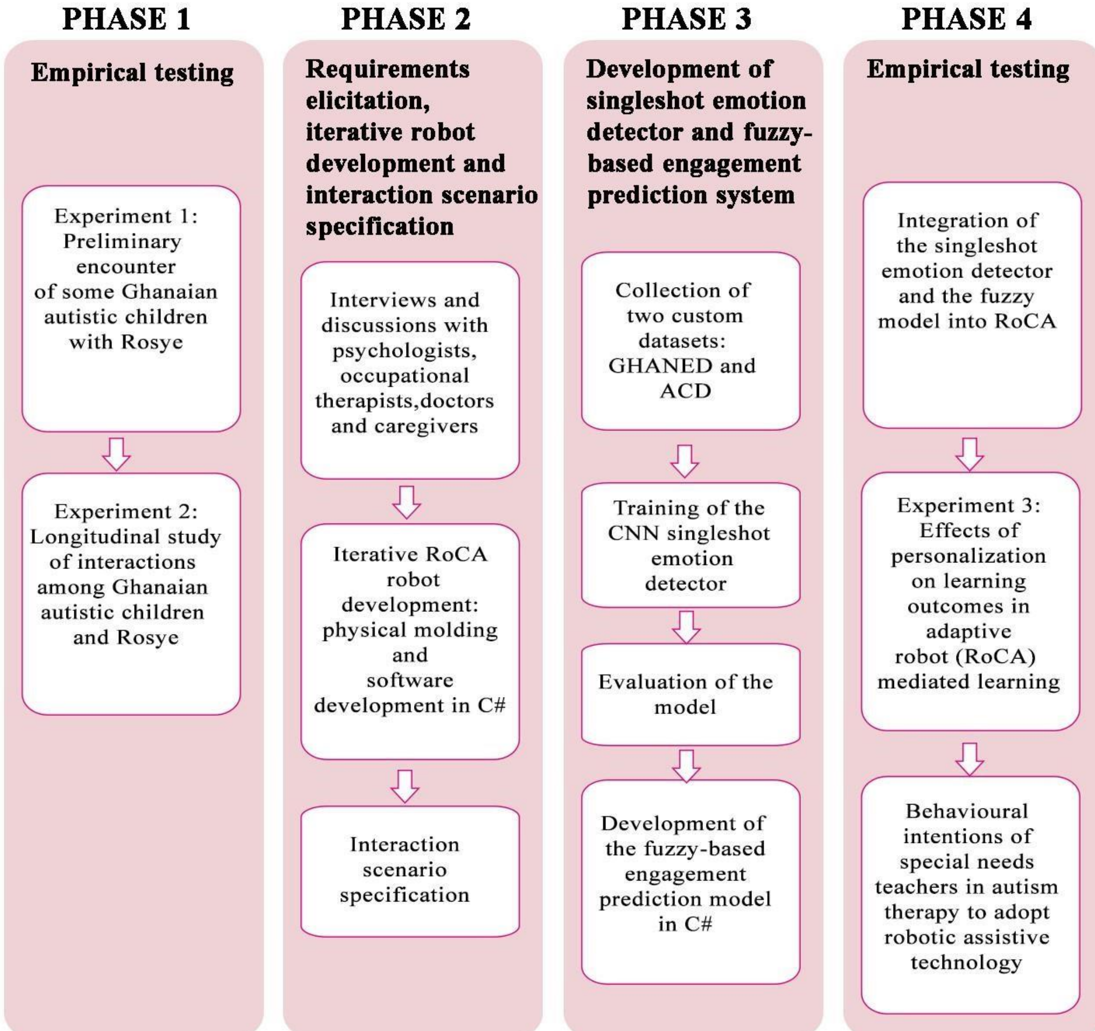


Figure 3.1 methodological structure of the research process

3.1 Preliminary and longitudinal empirical autistic child-robot interactions using Wizard of Oz (WOZ) approach

Two WOZ experiments were conducted in the first phase of the research.

- The first experiment investigated the effects of the uncanny valley phenomenon and the ability of the humanoid robot to engage the children in imitation games and general activities through multiple verbal prompts.
- The second experiment was a longitudinal study to observe the reaction and assess the engagement levels of the autistic children in their interaction sessions with a robot for a maximum of eight sessions on different days.

3.1.1 Experimental Research Approach

The approach used in this phase was experimental empirical research. Empirical research is a scientific method of investigation which enables an experimenter to acquire knowledge, gain direct experience, report on observations and present findings from a field experiment (Roth, 2007). It is suitable for research when specific research questions need to be addressed. Empirical research was used because the nature of research questions indicated that, field experiments needed to be conducted to be able to assess the effects of the uncanny valley phenomenon, impact of longitudinal interactions and effectiveness of the deep-fuzzy engagement estimation framework during autistic child – robot interactions. In autism management, there is a saying that" if you know one person with autism, you only know one person with autism". This adage implies that, no two people on the autism spectrum have the same set of characteristics and therefore, individualized therapies are better suited for them.

With this in mind, the experiments were conducted using the single subject design approach. Single subject design is a quantitative method which involves studying in details the behaviour of a small number of participants and each individual serves as his or her own control during the experiment (Alnahdi, 2013). This strategy enabled the individualization and assessment of the effects of robot mediated interventions on each individual. A "within-subject approach" was utilized to expose each participant to the two treatment conditions i.e. performing the imitation game and providing responses to tasks in the general activity session. A mixed method approach was utilized to collect both quantitative and qualitative data.

3.1.2 Case study

The experiments were conducted in two schools for children with autism in the Greater Accra region namely Autism Awareness Care and Training Center and HopeSetters Autism Center. Both centers are privately funded and rely on the fees paid by parents and the benevolence of philanthropists to run the institutions. In both schools, the children are taken through academics and social skills training, occupational, art and speech therapy as well as independent living skills. In one of the schools, each child is assigned a care giver, who guides the child through a unique curriculum developed for the child based on his or her abilities. In the other school, a group of children are assigned to a care giver who takes the children through various lessons. At the Autism Awareness Care and Training Center, the experiments were conducted in the sensory room. A sensory room is any room which has been filled with materials with varying stimuli such as soft balls, colours and lights to provide a safe environment for children with special needs to interact and explore with minimal risks. At HopeSetters Autism Center, the sessions took place at a general office.

3.1.3 Demographic information of the study participants

Invitation and consent forms were sent to parents of twenty (20) children with autism. Out of this number, fifteen parents agreed for their children to participate in the experiments. The inclusion criterion was children between the ages of 9 and 17, who had been previously diagnosed with autism. Fifteen children (n=15) who are on varying levels of autism spectrum participated in the preliminary experiment. Out of this number, three (3) were girls and twelve (12) were boys; mean age was 12.4 and standard deviation 2.47. Eight (8) of the children attended school at HopeSetters Autism Center and seven (7) of the children schooled at Autism Awareness Care and Training Center.

In the longitudinal studies, the participants were seven (7) children all of whom were from Autism Awareness Care and Training Center. Two (2) of them were girls and the remaining five (5) were boys. The mean age was 10.4 and standard deviation was 1.6. The number of participants in the empirical studies was minimal to be able to study the behaviour of the children in much details in correspondence with most single subject experiments usually called small-n designs. For the purposes of privacy and anonymity, the names of the participants have been pseudonymized.

3.1.4 Robotic platform

The adult sized humanoid robot, Rosye was used to conduct the WOZ preliminary and longitudinal studies. Rosye (figure 3.2) is a simplistic robot equipped with accessories such as ultrasonic sensor, led lights and speakers.

Figure 3.2 The humanoid robot Rosye

3.1.5 Experimental Script

The same experimental script was used in the preliminary and longitudinal studies. Each child-robot interaction session took place in either the sensory room or the general office depending on the Autism Center. The sensory room had many toys and playful objects which in a way were “distractions” to the experiment. However, in final deployment, the robot is likely to be situated in classrooms where there would also be other objects and humans. Therefore, a decision was made to determine whether the children would be able to focus on the robot despite the “distractions” in the room. The participants were accompanied by their caregivers; the presence of the caregivers provided a reassuring environment for the children. There was no fixed time for each child-robot interaction session. As a result, the children had the opportunity to interact with the robot as long as they felt comfortable and the sessions ended only when all the tasks had been completed. However, it was planned that a session would be brought to an end when a child became aggressive or felt uncomfortable.

The flowchart in figure 3.3 depicts the various stages involved in each child-robot interaction session. Each session followed a sequence of activities beginning with a quiet phase followed by an introduction, musical interlude, six imitation tasks, musical interlude, five general activity tasks and a concluding session. The robot was remotely controlled (WoZ) via a computer by the researcher and the instructions delivered by the robot were pre-recorded. Before any activity session, the robot informed the child of the tasks to be performed and at the end of the session, the robot also prompted the child. This principle adopted by the robot is in line with the protocols used in their classroom sessions where the children are prompted at the beginning and end of every activity session.

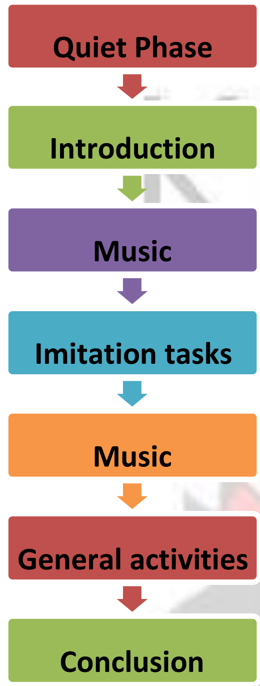


Figure 3.3 The seven stages involved in each child-robot interaction session

Stage 1: Quiet Phase

The robot was off as the child is ushered into the room in or der to observe whether the child would draw towards the robot, keep his or her distance or even express displeasure. The robot remained off for 30 seconds after which it was turned on by the researcher.

Stage 2: Introduction

The LED lights around the robot's eyes and mouth were turned on and the robot spoke for the first time. **Robot vocal(RV) 1**: "Hello, my name is Rosye. I am a robot".

Stage 3: Musical Interlude

At the third stage, the robot asked the child his or her name and told the child it wasgoing to sing a song.

RV 2: What is your name?

RV 3: Hello "*name of child*", I am going to sing a song for you

RV 4: Robot sang a popular local song and danced along using hand and neck movements

RV 5: Music time is finished

Stage 4: Imitation tasks

After playing the music, the robot led the child through physical exercise sessions. The sessions were meant to teach the children imitation skills. Robot raised its left, right and both hands up and down and encouraged the children to do same.

RV 6: It is exercise time!

RV 7: Left hand up
Action: Robot raised its left hand up

RV 8: Left hand down
Action: Robot put its left hand down

RV 9: Right hand up
Action: Robot raised its right hand up

RV 10: Right hand down
Action: Robot put its right hand down

RV 11: Both hands up

Action: Robot raised both hands up

RV 12: Both hands down
Action: Robot put both hands down

RV 13: Exercise time is finished

Stage 5: Musical Interlude

After the imitation tasks, robot played music while moving its hands and neck

RV 14: Robot sang another popular local song

RV 15: Music time is finished

Stage 6: General Activities (GA)

In the General Activity session (GA), a task was evaluated as completed when a child turned to look at the robot or the obj ect the robot is pointing at or performed the robot's request depending on the scenario. The robot was operated to deliver every instruction three times and if the child did not respond by the third prompt, task was aborted. The robot rewarded a child for good work done by saying "**good job**!"

GA task 1

Just before the activity session, each child was given a toy to play with. The robot tried to shift the attention of the child from the toy towards the robot. The robot mentioned the child's name.

RV 16: "*Hello name_of_child*"

Robot repeated **RV 16** till the child turned his or her attention towards the robot or aborted the task when the child did not turn by the third prompt.

GA task 2

Whereas some of the children have speech, others are nonverbal. The robo t asked the children the question “how are you” because almost all the children are able to respond to this question using speech or sign language.

RV 17: “Hello “*name of child”* how are you”?

GA task 3

Robot tried to shift the attention of the child towards a ball on the floor.

RV 18: (“name_of_child”), take the ball (while robot pointed to the ball)

GA task 4

Robot then asked the child to give the ball to it.

RV 19: Hi (“name_of_child”), give me the ball

GA task 5

Robot asks of a child’s name.

RV 20: What is your name?

Stage 7: Conclusion

Robot then informed the child the playtime and interaction session had ended.

RV 23: Hello my friend, we are done for today. You have done so well. Good job. Bye bye.

3.1.6 Data Collection and Analysis

The data collection strategy utilized in the preliminary and longitudinal experiments involved direct observations and video recordings. Due to the nature of the experiments, relying on only direct observations as the data collection technique would have posed challenges in recalling the behaviour of each child during the experiment at the data analysis stage. With the consent of the parents, all the interaction sessions were video recorded. Video recordings facilitated capturing of complex interactions simultaneously and provided access to multiple views of the recordings. Analysis of the recorded data was done by visual inspection of the video of each child-robot interaction session and subsequently, statistical methods were used to model the results. Each participant's data was analysed and plotted in order to make judgements about his or her response to the robot intervention on the first day, and for the participants in the longitudinal studies, their reaction to the robot over time. Each task had a score of one when correctly done by a child and a score of zero when task was not completed.

The specific metrics used in the preliminary experiment were:

i. prompt level at which task (imitation and general activity) was done: task done at first prompt, second prompt, third prompt or task not done.
ii. The number of times each child touched a specific part of the robot: eyes, neck, mouth, hand, shoulder, midsection and head

The metrics used during the data analysis of the longitudinal study were:

i. The imitation and general activity score per day for all the sessions (at most eight) ii. The frequency of physical contact with the robot measured over all the sessions iii. Comparison of the children's responses to both tasks (imitation and general activity) against the number of times each child touched the robot.

After the preliminary and longitudinal studies, a novel humanoid robot was developed. Sections 3.2 and 3.3 outline the robot design and development process.

3.2 Participatory design of RoCA, a humanoid robotic assistive technology for children with autism

More often, educational technological systems are developed without pedagogical consultations and planning with target users. Consequently, teachers are "forced" to adapt their teaching strategies towards the capabilities of developed systems. Although some recent studies have highlighted the potential suitability of robots in autism therapy, research gaps yet to be considered include selection of appropriate software, hardware and robot designs and development of relevant interaction scenarios. Research on robot enhanced therapy for children with autism is still in infancy as most of the robots developed are in the prototype stages. There is still room for vast consultation between artificial intelligence systems developers and stakeholders in autism therapy.

The robot, Rosye which was used in the preliminary and longitudinal study is an adult sized multi-coloured robot with limited functionality. Feedback from the two WOZ studies indicated that for long term autistic child-robot interaction studies, the physical embodiment of the robot needs to appeal to the children as a playmate rather than an adult teacher. More so, due to colour sensitivity in autism, the colour of the robot needed to be neutral to minimize sensory issues. Through series of consultations with the caregivers and stakeholders, it was decided that a smaller neutral coloured robot would be more appropriate for long term interactions considering the nature of the disorder.

To address this problem of robotic technology mismatch and end user expectations, a participatory design research approach was utilized to develop a social robot with and for children with autism. Through observations and interviews, the strengths and weaknesses of the children which could influence the design of the robot were identified. The knowledge gathered from the requirements analysis phase was incorporated into the design of the low cost robot RoCA which can be deployed especially in resource constrained environments which have pressing need for assistive technologies for children with autism.

3.2.1 Requirements Elicitation

In human-computer interaction, interviews, direct observations and surveys can be used as requirement elicitation techniques to understand the concerns, expectations, attitudes and needs of people who may interact with assistive technologies. The two elicitation techniques utilized to gather the requirements for the design of robot were direct observations and interviews. Some Autism Centers in Accra and the special school in Kumasi were visited on numerous occasions to familiarize with the children and their dayto-day classroom activities and play sessions. The visits also afforded the opportunity to observe the children and their teachers in their daily teaching sessions and to collect data which was useful in the design and construction of the robot. Interaction sessions were also held with some of the children who were verbal and could sustain conversations. Open-ended questions were presented to various stakeholders (teachers, psychologists, parents and other caregivers) in autism care during one-on-one interview sessions.

During the visits, it was realized that some of the children could exhibit unpredictable, aggressive and self-injurious behaviours. Others had sensory issues with touch, colour and light. These factors informed the choice of hardware and software resources for the robot development. The requirements elicitation phase ended with the establishment of functional specifications for the development of a child-sized humanoid robot RoCA. The robot needed to be equipped with motion capabilities, sound, camera, recording capabilities and adequate lighting.

Aside the functional requirements, non-functional requirements were specified as follows:

i. Appearance: The robot should provide a simplistic visual appeal to the children. The colours, lights and materials used in building the robot need to be user-friendly to the autistic children.

ii. Safety, Security and Usability: The software controlling the robot should be programmed to provide a safe and conducive environment for the children and their care givers. An authentication scheme should be provided to ensure authorized control of the robot. The interface for controlling the robot should be easy to use.

iii. Energy: The robot should be developed with electronic components which provide efficient use of power supply and generate very little heat.

3.2.2 Interaction scenarios specification

The choice of scenario media (text, images. videos or live interaction) could influence the responses of people towards robots (Xu et al., 2015). Robots which enact scenarios familiar to people are likely to gain acceptance as compared to those which portray arbitrary behaviours. A user-centered approach was adopted to engage teachers of the autistic children in the co-creation of activity scenarios which were used in the interaction sessions. Scenarios presented by the robot were streamlined to encompass activities from their learning sessions and their play sessions (imitation game).

3.3 Iterative robot redesign, development and prototyping

Based on the accrued requirements, the structure of RoCA was decided on as a naturalistic embodied humanoid because biologically inspired robots are usually appealing to humans and are able to blend into the environment quite easily. Appropriate hardware and software were selected to serve as building blocks for the robot development and control.

3.3.1 Hardware approach

In selecting materials for the construction of the robot, factors considered were: the properties of the materials (strength, toughness, and density), suitability, costs and availability. The materials chosen for the robot development are as follows:

i. Electronic components: The electronic components used were EZ-B v4/2, servos, camera, jumper cables, multicoloured LEDs, speaker and 16V batteries ii. Polystyrene: Polystyrene is a light-weight, recyclable and inexpensive foam with a closed-cell structure making it difficult for water to permeate. It also provides thermal insulation thereby reducing the amount of heat transferred among the various components of the robot.

iii. Lacquer: To achieve a good waterproof robot, lacquer was applied on the moulded polystyrene.

iv. Supergrip: Supergrip is a water-resistant contact adhesive glue that can bind together a wide variety of materials such as leather, polystyrene, paper and rubber.

It was used as a hardener and protective surface coating of the robot.

v. Plexiglas: Illuminating material used for the robot's face vi. Skimming Plaster: Applied over the robot to give it a smooth feel vii. Sol and Chuva: Rubbered paint from Coral and Dulux applied on the robot to give it the whitish colour *viii.* Magic dry erase: Final coating applied on the robot to provide a glossy and dryerase surface

3.3.2 Robot Construction method

RoCA is a full body 1210mm tall humanoid robot running on two 32 -bit ARM Cortex processors. It has been equipped with EZ -B v4/2 Wifi robot controller, 16V DC battery, loudspeakers, microphones, a wireless camera and customizable multicoloured LED lights to provide visual appeal to the children. The body of RoCA is made from polystyrene, a light weight, rigid but moldable material with the ability to maintain stability. The physical structure of RoCA, which is easy to assemble and disassemble, is composed of a humanlike head, a neck, upper body with two arms and lower body containing wheels and continuous rotation servos at the base to drive the robot. Each arm of the robot consists of a shoulder with 1 degree of freedom (dof). In consultations with experts, the robot has been designed to look simple in order to appeal to the children on the autism spectrum. The construction phase of the robot began with the specification of the conceptual view of RoCA shown in figure 3.4.

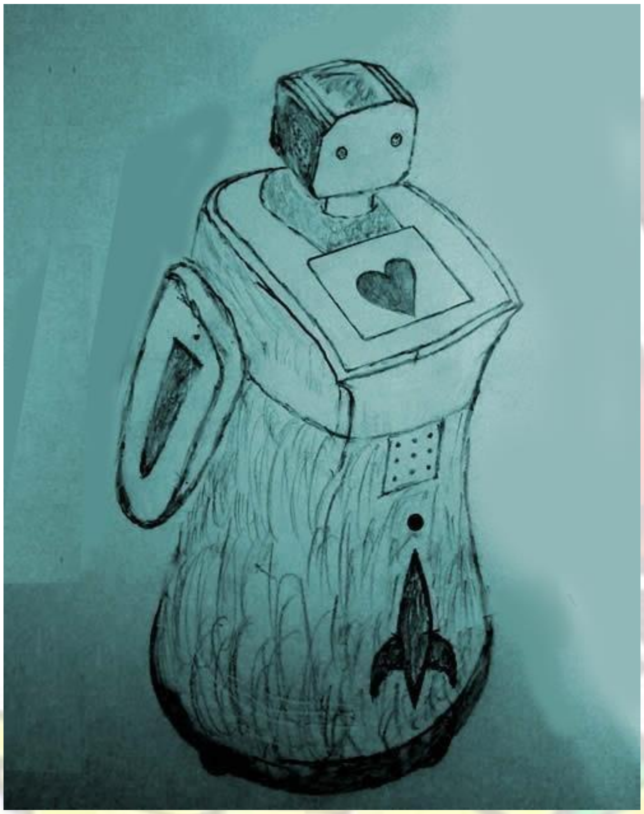


Figure 3.4 Conceptual drawing of RoCA

The construction approach used in the development of the robot was the “scratch -build” strategy. The robot body was carved out of the major raw material, polyst yrene foam according to the required dimensions. Polystyrene sheets were pieced together to achieve the required width and thickness and the sheets were cut into sections and separated; representing the different parts of the robot. The actual shapes weremolded starting from the base, upper body, chest, neck and head. Sanding was performed on the molded robot parts in order to achieve the required smoothness level. Hardening agents were applied to the outer surface of the molded parts to achieve the right sturdiness. Holes and compartments were created in the robot to allow cables to pass through and joints to fit together. Figure 3.5 depicts the structural molding process using polystyrene.

Figure 3.5 Structural molding process of RoCA

Surface finishing was completed by applying Sol and Chuva and magic dry erase. Molded parts were joined together to form one complete unit and the electrical components were fixed into the robot. Servo motors were mounted at the movable joints; neck, left and right arms and base. Cables were then carefully passed through the holes and compartment from all the joints to the main control board located at the base of the robot. All the electronic components were tested to ensure no breaks in the cabling.

3.3.3 Software Approach

The brain of RoCA is EZ-B v4/2, a wifi robot controller which can be manipulated via a computer or mobile phone. This robot controller is more powerful and faster than Arduino. With the EZ-B v4/2, most of the controlling programs reside on a computer instead of the memory onboard the microcontroller hence the robot can be programmed to execute complex tasks without worrying about storage limitations. EZ-B v4/2 has 3 I²C (InterIntegrated Circuit) ports, 3 UARTs (Universal Asynchronous Receiver/Transmitter), 24 multi-use servo/digital/serial ports and 8 analog ports. Two Integrated Development Environments (IDE), EZ-Builder and Visual Studio 2017 were used as platforms to program the EZ-B v4/2 using Ez-Script and C# respectively. The minimum system requirements needed to be able to install EZ-Builder are: Windows 8.1 minimum, Intel Pentium or AMD 64 or 32 Bit, 1.8 GHz CPU, 6 GB RAM and 200 MB free drive space.

3.4 Development of the Singleshot Emotion Detector (SED) based on transfer learning

After the robot development phase, the next research undertaken was the development of a novel deep learning based emotion detection system for autistic children. In this section, a description of the datasets, preprocessing strategies, training, testing and evaluation metrics for the proposed Singleshot emotion detection model are presented. In hardware systems especially for robots, there are technical constraints that have to be considered. Autonomous robots have limited computing power, limited onboard memory, low latency requirements and power efficiency requirements.

The proposed emotion detection approach, SingleShot Emotion detector (SED) was developed to classify emotions of autistic children based on images captured from a live video streamed by the robot's camera. Unlike traditional emotion recognition systems which separate face detection and classification pipelines making them less suitable for robotic applications due to their high latency, coupling and low frames per second(FPS), the proposed model combines face detection and classification into a single pipeline in order to provide a fast and reliable emotion prediction system. SED was implemented in Tensorflow, an open-source library written in Python and C++.

3.4.1 Architecture of the proposed SingleShot emotion detector

Object detection systems such as faster-RCNN and YOLO are computationally intensive (Liu, et al., 2016) for robots and would consequently be slow for real-time autistic childrobot interactions. SSD is a CNN known for its speed and accuracy in object detection (Liu, et al., 2016). The proposed emotion prediction model adopts transfer learning on SSDLite, a mobile friendly version of SSD. The architecture of SSD is a base network of stacked convolutions of decreasing size, followed by SSD layers (conv. 6, conv. 7, conv. 8, conv. 9, conv. 10, conv. 11) and prediction layers (Liu, et al., 2016). The main difference between SSD and SSDLite is that, all regular convolutions in SSD are replaced with depthwise convolutions followed by 1*1 projections in SSDLite architecture. For a normal convolution operation with input data of size Df * Df * M, where Df * Df is the size of the image and number of channels, M (3 for an RGB image). Suppose there are N filters/kernels of size Dk * Dk * M, a normal convolution performed with these parameters results in an output size of Dp * Dp * N. For the number of filters, N, each filter has to slide vertically and horizontally Dp times, hence the total number of multiplications for a regular convolution is N * Dp^2 * Dk^2 * M.

With depthwise separable convolutions, the process is broken down into two, namely, depth-wise convolutions and point-wise convolutions (Sandler et al., 2018). In depthwise convolutions, convolutions are applied to a single channel at a time unlike regular convolution operations which are applied to all M channels at once (Sandler et al., 2018). In depth-wise convolutions, the kernel size is Dk * Dk * 1. In instances when there are M channels, M filters are required hence the size of the output generated will be Dp * Dp * M. Dk * Dk multiplications will be performed for a single convolution operation. The filter would have to be slided Dp * Dp times across all the M channels, hence the total number of multiplications will be M * Dk^2 *Dp^2.

In the second stage of depthwise separable convolutions, pointwise convolutions are performed. Pointwise convolutions apply 1*1 convolution operations on the M channels using a filter size of 1* 1 * M. If the number of filters are N, the output size becomes Dp * Dp *N. A single convolution operation would require 1*M multiplications with filter being slided Dp*Dp times. The total number of multiplications for a pointwise operation will be M * Dp^2 * N. The total number of multiplications for a depthwise separable convolution is the sum of depth wise convolution multiplications and point wise convolution multiplications (M * Dk^2 * Dp^2 + M * Dp^2 * N). This means that, depthwise separable convolutions are computationally less expensive and uses fewer parameters compared to regular convolutions, hence making them more suitable for deployment on mobile devices, robots and embedded systems. This is why the proposed Singleshot emotion detector (SED) adopts transfer learning on SSDLite which uses depthwise separable convolutions instead of regular convolutions.

SSDLite uses MobileNetV2 as feature extractor. MobileNetV2 contains an initial connected layer with 32 filters, followed by 19 inverted residual bottlenecks (for memory efficient computations), Relu6 activation function and a kernel size of 3x3 (Sandler et al., 2018). The inverted residual with linear bottleneck layers take as an input a lowdimensional compressed representation which is first expanded to high dimension and filtered with a lightweight depthwise convolution (Sandler et al., 2018). The features are subsequently projected back to a low-dimensional representation with a linear convolution (Sandler et al., 2018). This architecture adopted by MobileNetV2 makes it suitable for mobile applications because it reduces significantly the amount of memory needed to perform operations (Sandler et al., 2018).

The proposed model, SED adopts transfer learning on SSDLite model to two custom datasets generated in the course of the research. Transfer learning was used to facilitate the training process since pre-trained weights of the SSDLite model have learnt salient features in objects and therefore, minimizing the objective loss function with respect to the new dataset would take less time and less data to reach a desired accuracy. The first layers of most CNNs extract low level features common to many tasks. Therefore, all the layers except the final layer of SSDLite were frozen and the final layer was retrained using the custom datasets. A two stage fine-tuning approach was adopted to first of all, train the model on the GHANED dataset consisting of 532 images of some Ghanaians expressing both "acted" and "in the wild" emotions followed by training on a custom dataset of autistic children, ACD consisting of 287 facial expressions of autistic children collected during the preliminary and longitudinal ACRI sessions. According to Ng et al. (2015), supervised fine tuning on relevant facial expression datasets before training on small target datasets lead to high accuracy.

Figure 3.6 depicts the development process undertaken to derive the proposed system.

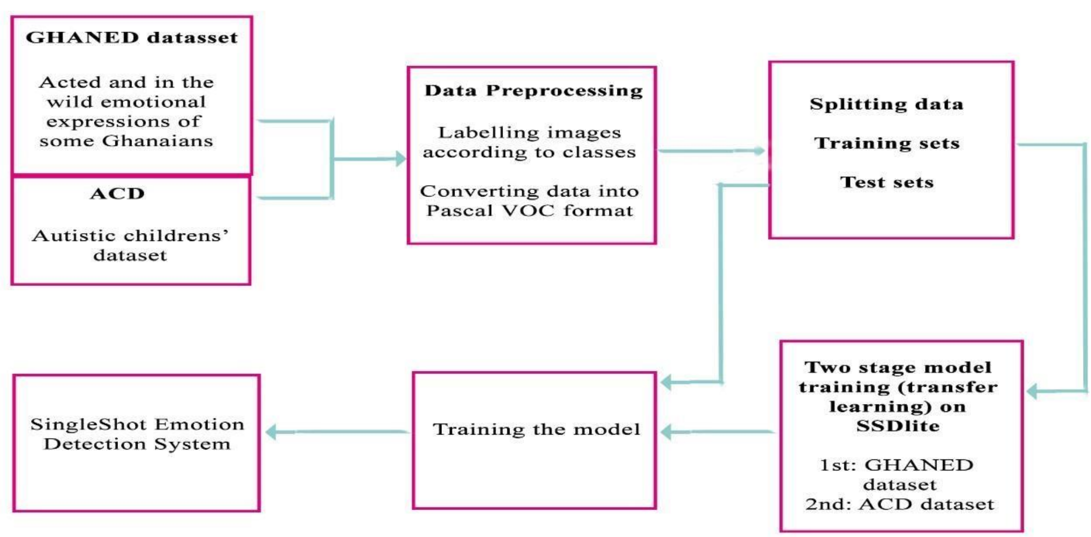


Figure 3.6 Development process of the proposed SingleShot Emotion Detector

3.4.2 Datasets

Emotion recognition systems sometimes face challenges in multicultural scenarios since emotions are likely to vary by culture (Quiros-Ramirez & Onisawa, 2015) or may lose some meaning across cultures (Elfenbein & Ambady, 2002). Existing FER systems have not been trained on facial expressions of children with autism and therefore those models would perform poorly in autism domain specific applications. A goal of the research was to develop a FER system for the humanoid robot, RoCA for use in robot mediated therapy for Ghanaian children on the autism spectrum. As such, two datasets GHANED and autistic children's dataset (ACD) were collected. GHANED dataset contains images of some Ghanaians depicting six emotional classes: happy, sad, anger, fear, neutral and surprise. ACD dataset contains facial expressions of autistic children which were collected during the preliminary experiments and the longitudinal studies. Table 3.1 and 3.2 present a summary of the datasets and the number of labels per class. Sample images from the GHANED dataset are shown in figure 3.7.

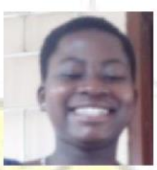
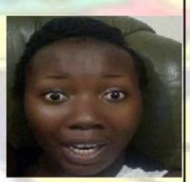
Table 3.1 Overview of the GHANED and ACD datasets

	Number of labels						
Dataset	Happy	Sad	Fear	Surprise	Neutral	Anger	Total
GHANED	92	87	88	90	89	86	532
ACD	50	47	43	49	50	48	287

For each dataset, 80% of the images were used as training samples and the remaining 20% were used as test sets. Splitting of the datasets allows the evaluation of the machine learning model on new data, that is data the model has not seen before. Also, the approach of evaluating the model on test sets helps to generalize the performance of the model with respect to unseen data. Sample images from the ACD dataset are not included due to privacy reasons.

Table 3.2 Overview of training and test sets

Dataset	Training set	Test set
GHANED	426	106
ACD	230	57

Happy	Sad	Surprise	Fear	Anger	Neutral
					

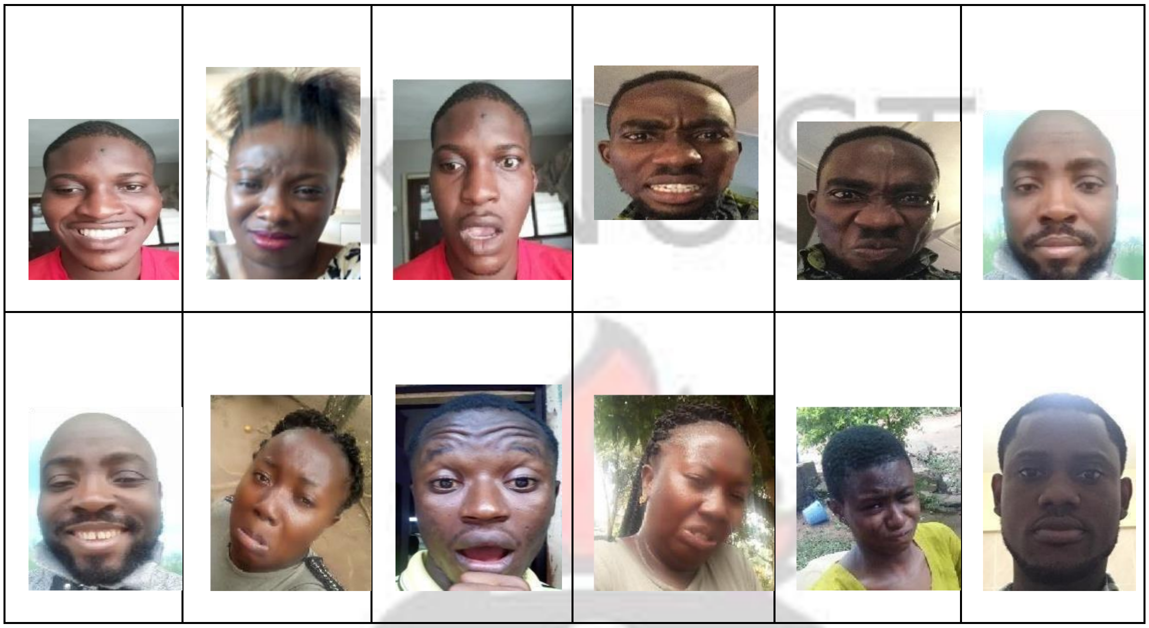


Figure 3.7 Sample images from the GHANED dataset

3.4.3 Preprocessing of facial expressions data

The training images were annotated in Pascal Visual Object Classes(VOC) format. In the Pascal VOC (figure 3.8), the annotation consists of the training folder name, file name, path to the file, bounding box coordinates and size of the image (width, height and depth) compiled in an XML file.

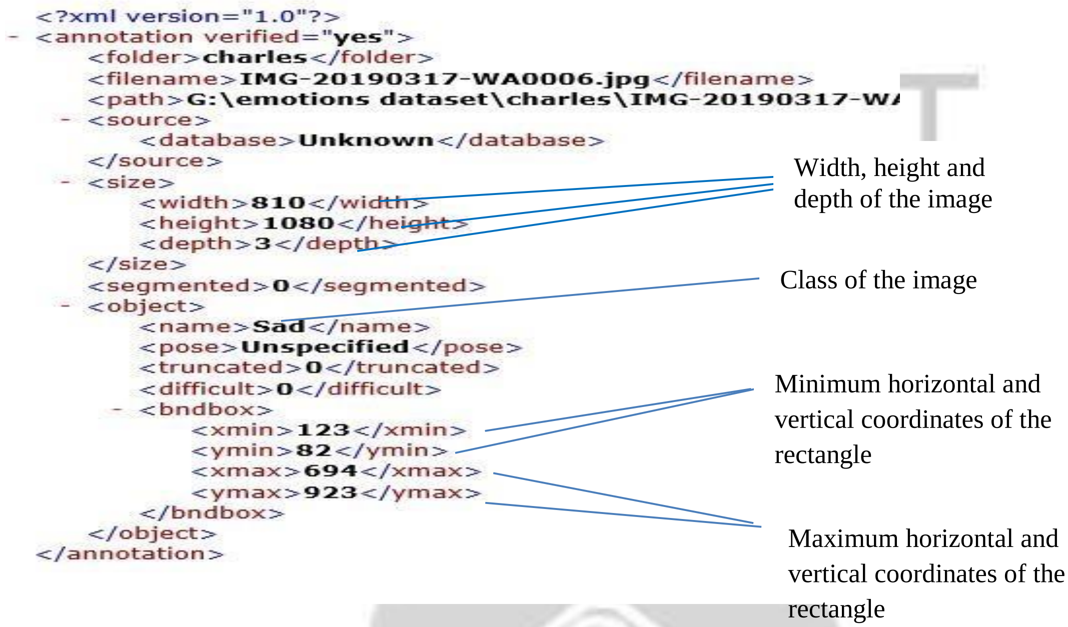


Figure 3.8 Sample XML file of containing an image labelled according to the Pascal VOC format

3.4.4 Training

The training of the images was done on a Tesla-K80 cloud-based Graphics Processing Unit(GPU) with Jupyter Notebook installed. The images together with their Pascal VOC annotations were converted into Tensorflow records. Tensorflow records are datasets converted into binary format in order to reduce the size of datasets so as to increase the speed of the training process. The tensorflow records were then fed into a training script provided by the object detection API. The training process was run for 4700 epochs with a batch size of 8. Random flip and random crop data augmentation techniques were applied to the data during the training phase. Tensorboard was used to visualize the training process and monitor performance of the model over the epochs.

3.4.5 Performance evaluation of the SingleShot Emotion Detector

The proposed model was evaluated using mean Average Precision (mAP), a common metric for evaluating object detection algorithms. Average Precision is based on precision and recall where precision measures the "*false positive rate" or the ratio of true object detections to the total number of objects that the classifier predicted*" and recall measures the "*"false negative rate" or the ratio of true object detections to the total number of objects in the data set*" (Arlen, 2018). Mean Average Precision or mAP score is calculated by taking the mean of the interpolated average precisions over all classes (Henderson & Ferrari, 2016). The model was also evaluated in a human-robot interaction scenario where the robot RoCA was programmed to readapt teaching sessions based on the emotional states and engagement levels of autistic children.

3.5 Fuzzy-based engagement prediction framework for robot-mediated learning for children with ASD

The first two empirical studies were conducted via the Wizard-of-Oz approach where the robot was controlled unbeknownst to the children. Feedback obtained from the children and their caregivers indicated that, although the children reacted well with the robot, further improvements could be made. In a multi-diversified area such as autism management, social robots need to be able to act as interaction agents which can personalize their behaviour based on individual profiles, affective states and learning progress. One key strategy which was identified was to equip the robot with some level of autonomy to be able to detect the affective states and engagement levels of the autistic children during real-time interactions in order to readjust the learning process accordingly.

To this end, a personalized interactive framework (figure 3.9) which utilizes the SingleShot motion Detector (SED) model and a fuzzy inference system (FIS) to foster sustained and meaningful long -term autistic child -robot interaction sessions has been proposed. This framework was illustrated and assessed by integrating the SED model and the FIS system programmed in C# as a plugin into RoCA in order to assess the effects of personalization on learning gains.

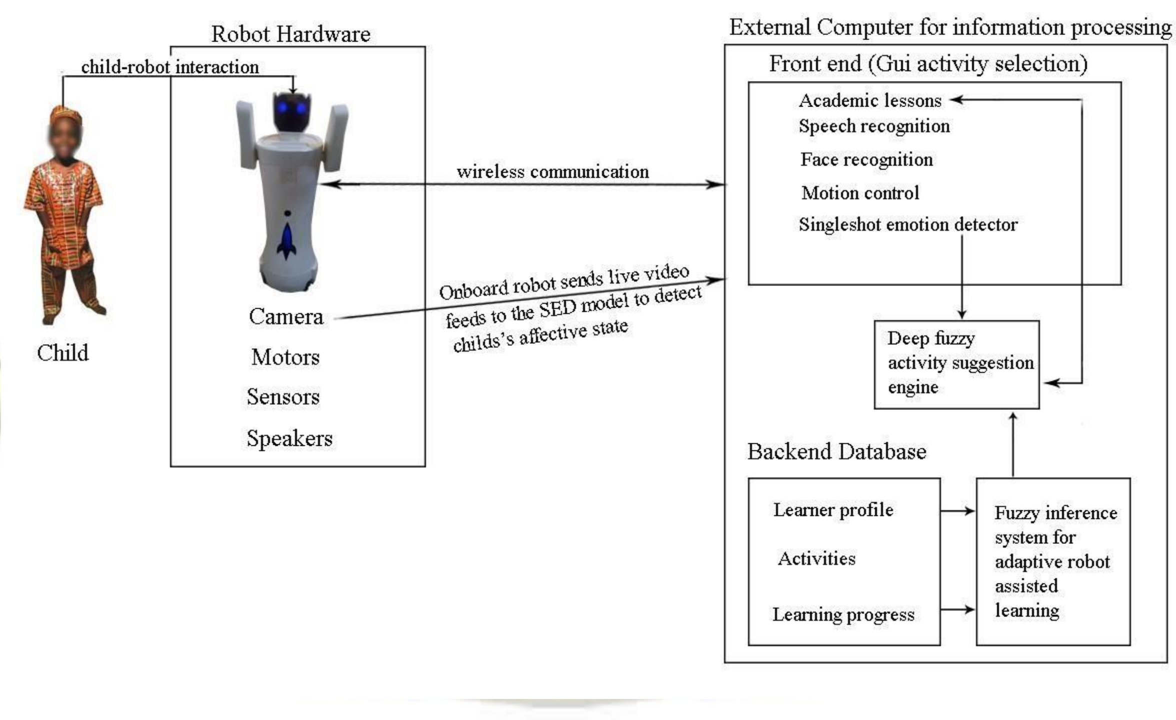


Figure 3.9 The proposed deep fuzzy robot behaviour adaptation framework

The SingleShot Emotion Detector (SED) presented in section 3.4 was modelled as a plugin using Visual Studio 2017 and then it was integrated into RoCA's custom software interface control which was programmed with Ez-Builder. To be able to stream live videos from the robot's camera to the SED model, the camera device and HTTP Server were started in Ez -Builder. The HTTP Server makes sure that applications can access the robot's camera via a URL. The URL of the robot's camera was configured in the deep learning model and live image streams were sent from the robot to the SED model using OpenCV. Figure 3.10 indicates a snapshot of the codes responsible for capturing live feeds from the robot's camera.

```
ezb = ("http://192.168.1.2:80/CameraImage.jpg?c=Camera")
# Detection
with detection_graph.as_default():
    with tf.Session(graph=detection_graph) as sess:
        while True:
            sock = urllib.request.urlretrieve(ezb)
            capture = cv2.VideoCapture(sock[0])
            if (capture.isOpened() == False):
                print("Error opening video stream or file")
            # Read frame from camera
            ret, image_np = capture.read()
```

Figure 3.10 codes which capture live feeds from the robot's camera

After the frames from a live video feed were taken by the camera onboard the robot and fed into the SED model, the model performs a singleshot object localization and classification of the facial expression in the received input images. The inferred emotional classes of the localized faces are then sent to a file which logs the predicted emotional classes.

The fuzzy logic reasoning model, a component of the proposed framework, is an artificial intelligence enabled algorithm which equips the robot with the ability to reason and react appropriately to individual child's needs. Fuzzy logic is a robust approach for dealing with uncertainties and deriving conclusions based on imprecise, ambiguous and noisy data. The proposed fuzzy robot behaviour adaptation framework shown in figure 3.9 consists of three linguistic input variables and one output variable, where each variable has a set of membership functions derived from triangular membership function equation.

The inference engine of the fuzzy controller consists of 27 rules for measuring engagement which were derived together with some caregivers of children with autism. Output from the fuzzy inference system is combined with the highest aggregated predicted emotional class and the deep fuzzy activity selection engine suggests to the teacher the appropriate action to take if the child's predicted engagement is low or average. The proposed fuzzy based engagement prediction inference system was implemented in C# and is composed of four components namely: fuzzifier, knowledge base, inference engine and defuzzifier.

3.5.1 Determination of crisp input variables and crisp output variables

The fuzzy-based framework has three crisp inputs (linguistic variables) i.e. the score of a child at a specified time point in the learning session, task difficulty and IQ level. The engagement level of the child is the crisp output variable. Figure 3.11 shows the interface of the fuzzy controller integrated into the control system of RoCA, where the linguistic variables and their fuzzy values are set by the domain experts (teachers).

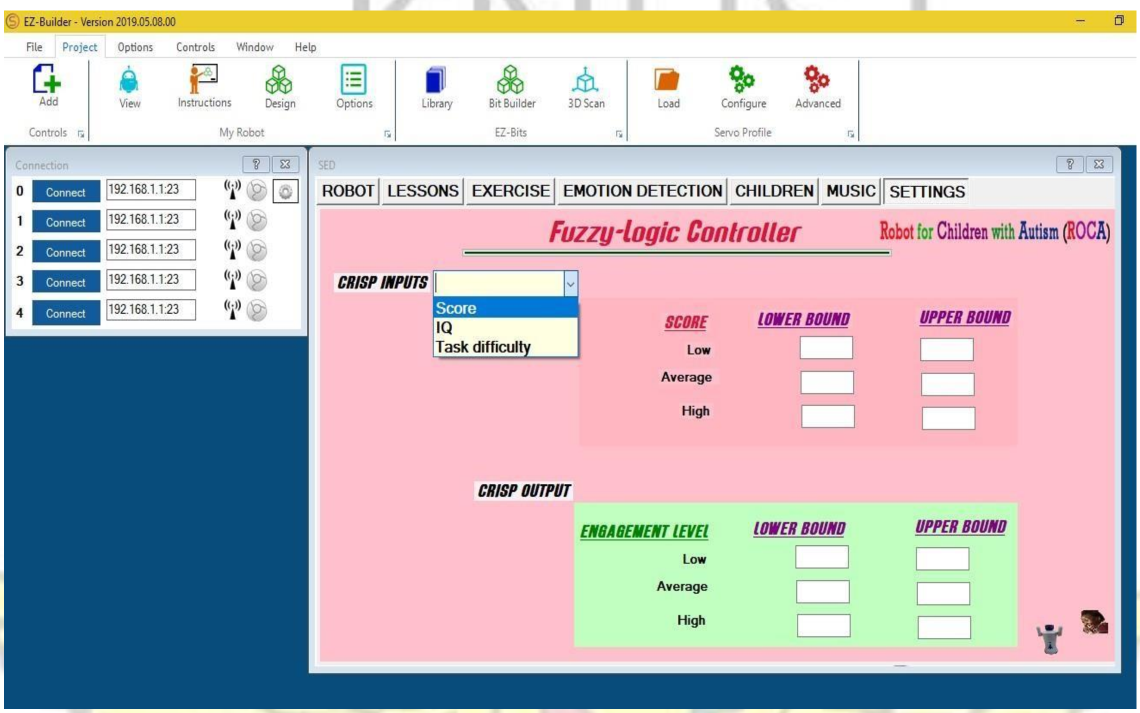
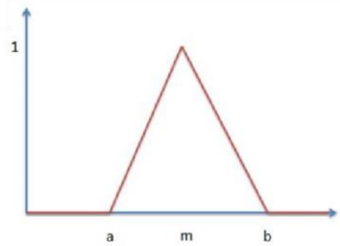


Figure 3.11 Fuzzy-logic controller settings for linguistic variables and fuzzy values

3.5.2 Fuzzification

In the fuzzification stage, fuzzy sets were constructed for the input and output variables. Although there are many other fuzzi fiers such as trapezoid, singleton and Gaussian, triangular membership functions were used since they are easily to transfer to microcontrollers (Moslehi, 2011). A triangular membership function can be represented as shown in figure 3.12.



$$\mu_A(x) = \begin{cases} 0, & x \leq 0 \\ \dfrac{x-a}{m-a}, & a < x \leq m \\ \dfrac{b-x}{b-m}, & m < x < b \\ 0, & x \geq b \end{cases}$$

Figure 3.12 Triangular membership Function and equation

The values of a, b and m were set by the teachers of the children.

All the input variables were fuzzified into triangular member functions using the equations in figure 3.12. The score of a child was fuzzified into 3 membership functions using linguistic terms: high, average and low. Task difficulty was fuzzified into member functions: easy, medium and hard. The IQ level was given membership functions low, medium, high. The engagement level of the child was fuzzified into three membership functions: highly engaged (HE), engaged (E) and not engaged (NE). The fuzzy rules were developed based on interactions with domain experts in care for children with autism. Table 3.3 shows sample rules for predicting engagement based on IQ, task difficulty and score.

Table 3.3 Sample rules for predicting engagement based on IQ, task difficulty and score

Rule number	Score	IQ	Task Difficulty	Estimated engagement level
1	Low	Low	easy	not engaged
2	Low	Average	easy	not engaged
3	High	Average	hard	highly engaged
4	Average	High	easy	Engaged
5	High	High	high	Highly engaged

3.5.3 Rule evaluation and Defuzzification

Mamdani inference approach was used to map the three input variables (IQ, score and task difficulty) into one output variable (engagement) using IF … THEN rules. For example, If score= low and IQ=low and task difficulty=easy THEN "not engaged". The output from the rule evaluation is always a fuzzy set. In the defuzzification step, the output was converted into a crisp value for use by the robot. Defuzzification techniques include center of gravity, center of sums and max criterion. Center of gravity was chosen as the defuzzification technique because it is computationally efficient (Djam et al., 2011) and most prevalent of all the other approaches (Owoseni & Ogundahunsi, 2016).

3.6 Evaluation of the deep fuzzy robot behaviour adaptation framework

An empirical study was conducted using two different setups of RoCA robot; RSEDFuzzy setup where the robot equipped with the SED model and the fuzzy inference engine operated semi-autonomously and the RWOZ setup where the robot was operated manually. In the RSEDFuzzy setup, the robot used factors such as the highest predicted emotional state, scores and task difficulty to deliver personalized lessons to children. For example, if the emotional states logged into the SED model indicates that a child is sad and the score of the child at a specified point in time is low but the task difficulty level is easy, it means that the robot needs to take an action (for example, play music or change the lesson) which can help re-engage the child. On the other hand, when the robot run using the RWOZ mode, the robot did not take into account any factors into consideration but just went on delivering the lessons till the sessions were over.

The aim of the empirical research was to assess the effects of RSEDFuzzy based personalization on learning gains in autistic children as compared to RWOZ random operated robot (a robot operated manually using wizard-of-oz). The research question that had to be addressed was: *Would the scores of the ASD children in a "fruit learning lesson" be higher in teaching sessions with RSEDFuzzy based personalized robot than a wizard of oz operated robot?* The independent variables were: RSEDFuzzy and RWOZ and the dependent variable was the scores for the tasks.

Twelve autistic people (11 males and 1 female) aged eleven to twenty-six who have been diagnosed with ASD and are enrolled at Garden City Special School in Kumasi participated in the study. A randomized controlled research design was adopted in the study; Six participants were randomly assigned to RSEDFuzzy group and the other six were assigned to the RWOZ group. The mean age in the RSEDFuzzy group was 18.0 and standard deviation was 5.4. The mean age in the RWOZ group was 16.7 and the standard deviation was 5.5. A Mann-Whitney U test was performed to test for significance in age difference between the RSEDFuzzy group and RWOZ groups. The results obtained, $U=16$ and $P=0.747$ indicated that there was no significant difference between the ages of the children in the two groups.

3.7 Experimental Setup

The experiments were conducted in the computer laboratory of the Garden City Special School using the humanoid robot RoCA. Each child had nine interaction sessions with the robot, resulting in a total of 108 interaction sessions. For each group, the study began with two pretest lessons (same for each group), five training sessions and two immediate posttests. All the children were accompanied to the experimental room by a teacher. The robot provided rewards by saying “good job” and used multiple verbal prompts when necessary. A tablet was fixed on RoCA to play video recordings of selected fruits. Two tables were placed beside RoCA with banana on top of one and apple on the other.

The first two sessions were pre-test lessons to test the knowledge of the children on identification of two fruits, apple and banana. Before each session, the name of the child was entered in a textbox on the robot control system. RoCA introduced itself by saying “*Hello + name of child” my name is RoCA. We are going to study fruits today. “name of child, take the apple on the table*” If the child picked the correct fruit, the child was given a score and the robot asked the child to give the fruit to his or her teacher. The same procedure was repeated for the other fruit i.e. banana. Anytime the robot asked a question, it waited for the child to be scored (by clicking good job if the task was done correctly). If the good job button was not clicked after ten seconds, the robot repeated the instruction again and waited for the “good job button click” or another fifteen seconds to deliver the final prompt after which it moved on to another question. After the question and answer session, the robot played and danced to a local song after which it told the child the session was over.

In sessions 3 to 7, RoCA used video modelling techniques and interaction sessions to teach the children to identify apple and banana and showed them how to eat the fruits. The lessons were restricted to two fruits in order to avoid sensory overload which is typical in autistic children. The sessions always began with RoCA mentioning the child's name and informing them of the lesson. E.g. *Hello +name of child, today we are going to learn fruits.* A picture of the specific fruit e.g. apple was displayed on the tablet fixed on the robot. RoCA mentioned the name of the fruit and a short video of the fruit was shown on the tablet. RoCA then pointed and turned head to the direction of the specific fruit lying on the table and mentioned the name of the fruit again. RoCA instructed "*name of child" take "name of fruit*". Rewards or multiple verbal prompts were used when necessary. The same procedure was repeated for the other fruit. Finally, the robot sang a song and ended the interaction session.

In the RWOZ group, the robot operated in a non-adaptive mode where it delivered learning sessions without considering the emotional state, child's learning progress and task difficulty level. In the RSEDFuzzy group, RoCA acted as an adaptive agent which personalized learning sessions based on the inputs supplied to the deep-learning model and fuzzy-based engagement prediction engine (aggregated emotional state, learning progress, IQ and task difficulty).

The last two sessions were conducted as post-tests to assess the children on the knowledge gained after the teaching sessions. The script delivered by RoCA in these two sessions were similar to the script used in sessions 1 and 2. In the pre-test and post test lessons, each task had a score of 5 when performed correctly by the child. Figure 3.13 shows the robot control interface used to select and allocate marks for each child.

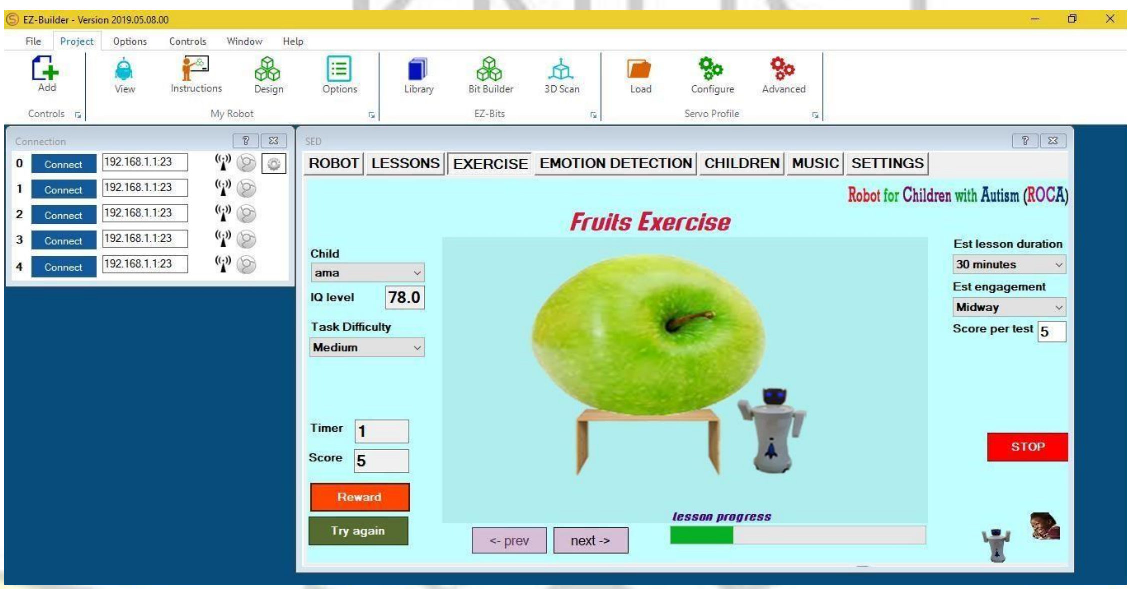


Figure 3.13 Control interface for pre-test and post-test lesson on fruits

3.7.1 Data analysis

The pre-test and post-test results of this study were analysed in SPSS. A Mann Whitney test was conducted to test for significant difference between the pre-test scores for both groups. The same test was used to examine the significant difference in the posttest scores for both groups. Mann Whitney test was used because it is a non-parametric and distribution free test suitable for small sample sizes and small test scores. The results of this study are presented in chapter 4 section 4.5.

3.8 Behavioural intention of special needs teachers to adopt robots as assistive technology in the classroom

Towards the end of the entire research, a survey was conducted to investigate the behavioural intention of teachers to adopt robotic assistive technology in special needs education for children with autism. This survey was based on the UTAUT model proposed by Venkatesh et al (2003) and aimed to study the impact of performance expectancy(PE), effort expectancy(EE), social influence(SI) and facilitating conditions(FC) on the behavioural intention of special needs teachers in autism therapy to adopt robots in the classroom.

3.8.1 Hypotheses

The hypotheses tested were follows:

H1: Performance expectancy will positively influence caregivers' behavioural intention to accept social robots in special needs education

H2: Effort expectancy will positively influence the behavioural intention of caregivers to accept social robots in special needs education

H3: Social influence will positively affect the behavioral intention of caregivers to accept robots in special needs education

H4: Facilitating conditions will positively affect the behavioral intention of caregivers to accept robots in special needs education

3.8.2 Instruments

The data collection method utilized in this experiment was survey. Questionnaires were developed based on the instruments proposed by Venkatesh et al. (2003) in the Unified Theory of Acceptance and Use of Technology (UTAUT) framework, to understand the underlying factors which would influence special needs teachers' willingness to accept robots as assistive technologies. Four independent variables performance expectancy(PE), effort expectancy(EE), social influence(SI) and facilitating conditions(FC) were identified. The dependent variable was behavioural intention (BI).

The participants had to indicate their level of agreement to twenty-three research statements in the questionnaire on a five-point Likert scale: totally disagree (1), disagree (2), undecided (3), agree (4) and totally agree (5). These questionnaires were administered to fifty (50) caregivers of children with autism. The questionnaire was sectioned into two parts: the first collected information on the demographics of the participants and the second part elicited information on the relationships among the dependent and independent variables. There were four questions relating to performance expectancy, four questions relating to effort expectancy, three questions directed at social influence, three questions for facilitating conditions and one question targeting behavioural intention. The questions asked in the aforementioned categories are as follows:

Performance Expectancy (PE)

PE 1: Robots would be useful in special needs education

PE 2: Robots would help accomplish teaching tasks more quickly
PE 3: Robots can increase productivity during school hours

PE 4: Robots will make it easier for me to do my job

Effort Expectancy (EE)

EE 1: I would find a robot easy to operate

EE 2: It would be easy for me to learn how to use a robot

EE 3: Using a robot in my work environment would be easy

EE 4: My teaching sessions using the robot would be clear and understandable

Social Influence (SI)

SI 1: My working environment support the use of robots

SI 2: I would use the robot if people who are important to me think that I should use robot for teaching.

SI 3: I would use robots if peo ple who influence my behaviour think that I should use robots

Facilitating Conditions (FC)

FC 1: I have the required knowledge to operate a robot

FC 2: Operation costs will encourage the use of robots

FC 3: A specific person should be available for assistance with robot operation difficulties

Behavioural intention (BI)

I intend to use robots in special needs education in future

3.8.3 Data analysis

A reliability analysis was conducted for each item in every construct i.e. PE, EE, SI, and FC using Cronbach's alpha in order to ascertain consistency in the survey results. Multiple regression analysis was performed in SPSS software to examine the relationship among the UTAUT constructs. The results of this survey are presented in section 4.6.

3.9 Informed consent

Consent forms were given to the three Schools as well as the parents of the children involved in the study. All the processes involved in the experiment were explained to them and they gave their consent for the research. Children with autism are classified under protected groups and it is important to maintain confidentially of the data collected. Therefore, the names of all the participants have been pseudonymized in this thesis.

In this chapter, the participatory design process for the robot design and development involving the selection of appropriate software and hardware platforms have been presented. The various steps undertaken in conducting the empirical studies, the singleshot emotion detector and fuzzy inference system have also been described. In chapter four, the results of the preliminary and longitudinal studies as well as the implementation details of the humanoid robot RoCA, results and evaluation of the emotion detector are presented.

CHAPTER 4

IMPLEMENTATION

This chapter reports on the results of the research activities undertaken. Two wizard-ofoz experiments were conducted to assess the initial reaction of some autistic children to a humanoid robot and their reaction to the robot over an extended period. Afterwards, a participatory design research strategy was used to design and develop a humanoid robot RoCA and another empirical study was conducted to assess the effects of personalization on robot-mediated learning gains. A survey was also conducted to investigate the behavioural intention of special needs teachers to adopt robots in the classroom setting.

4.1 Preliminary observations from interactions among Ghanaian autistic children and Rosye, a humanoid robotic assistive technology

The preliminary experiment investigated the effects of the uncanny valley phenomenon on the initial autistic child-robot interaction sessions and the ability of a humanoid robot Rosye to engage the children in imitation games and general activities. Fifteen (15) autistic children from two Autism Centers participated in this single day experiment. A summary of information on the participants are presented in table 4.1. The names of the participants have been pseudonymized due to privacy reasons.

Table 4.1 Gender, age and speech capabilities of participants

	Child ID	Sex	Age	Speech
1.	Afia	Female	9	Verbal
2.	Ama		12	Verbal
3.	Vida		15	Verbal
4.	Akwasi	Male	10	Verbal
5.	James		9	Verbal
6.	John		11	Verbal
7.	Yaw		9	Nonverbal
8.	Kofi		13	Verbal
9.	Jeff		17	Nonverbal
10.	Osei		13	Nonverbal
11.	Peter		13	Verbal
12.	Sarfo		13	Nonverbal
13.	Ike		16	Verbal
14.	Luke		13	Verbal
15.	Paul		13	Verbal

In the preliminary study, the selected participants had one-on-one interactions with the humanoid robot. The results from this study served as a baseline for comparing the initial responses of the children to their reaction to the robot over an extended period. People on the autism spectrum have different needs and preferences, hence it was expected that they would not exhibit the same behaviour towards the robot. One hundred and sixteen minutes forty-four seconds (116m 44s) of video recorded during the preliminary trial were analysed. The shortest child-robot session lasted for 2 minutes 53 seconds, maximum duration was 16 minutes 27 seconds and the average interaction time was 8 minutes 16 seconds.

4.1.1 Quantitative results from the imitation game

In some cases, the robot had to repeat th e instructions multiple times before the children responded. Figures 4.1 and 4.2 presents the scores of the children per prompt levels.

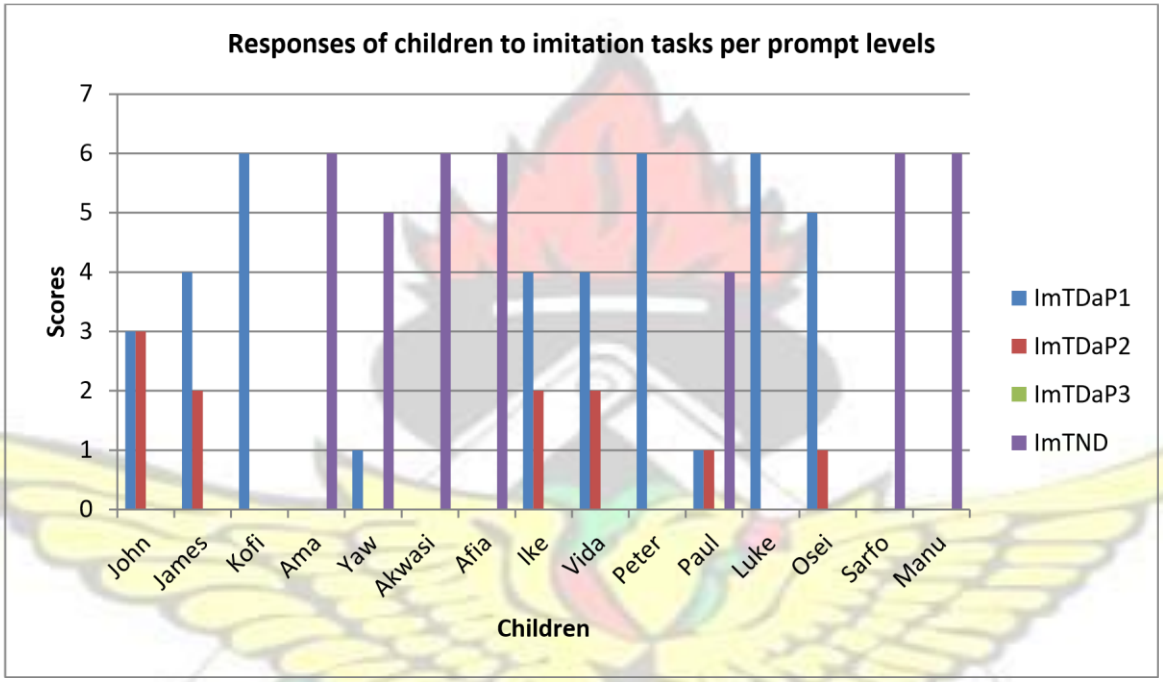


Figure 4.1 Responses of children to imitation tasks per prompt levels

ImTDaP1: Imitation tasks done at first prompt

ImTDaP2: Imitation tasks done at second prompt

ImTDaP3: Imitation tasks done at third prompt

ImTND: Imitation task not done

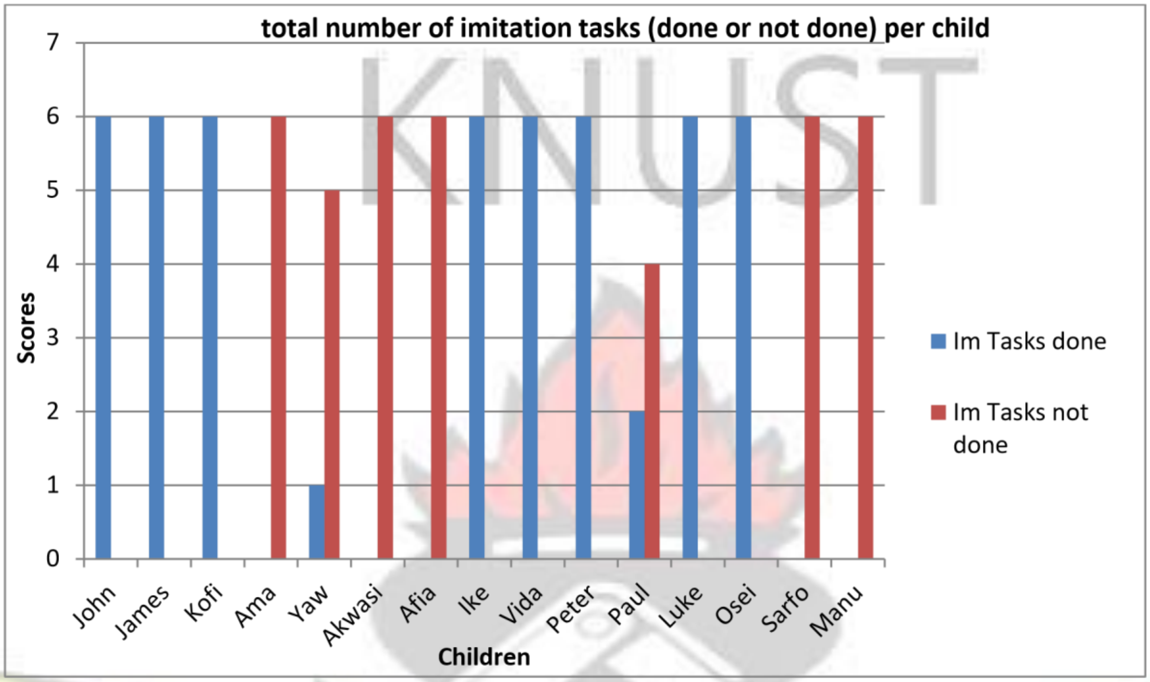


Figure 4.2 total number of imitation tasks (done and not done) per child

Im: Imitation

From figure 4.2, it is observed that 8 out of the 15 autistic children successfully imitated all the 6 actions of the robot during the imitation game. 5 children did not imitate any of the robot's actions. 1 child imitated only two of the robot's actions and 1 child imitated just one action. In all, the robot engaged in 90 imitation tasks with all the children (15 children and six imitation games for each child). Out of the 90 tasks, 40 were completed by the children at the first prompt given by the robot, 11 tasks were completed at the second prompt and the remaining 39 of the tasks were not performed at all. In figure 4.1, it is observed that, all the children who performed the imitation tasks did so at either the first or second prompts; none of the children performed an imitation task at the robot's third prompt. In all, 56.7% of the imitation tasks were successfully completed.

4.1.2 Quantitative results from the general activity game

Based on the personality and mood, some children on the autism spectrum may not engage in communication or respond to requests from their parents, caregivers and to a more severe extent unfamiliar people. Rosye was programmed to deliver five individualized instructions to each child (by referring to each child by name). Out of the 75 general activity tasks, 35 of them were successfully performed by the children at the robot's first prompt, 4 tasks were completed at the second prompt and 8 tasks were completed at the third prompt, unlike the imitation game where no tasks were performed at the third prompt. The number of GA tasks which were not performed was 28. Figures 4.3 and 4.4 present results from the general activity sessions.

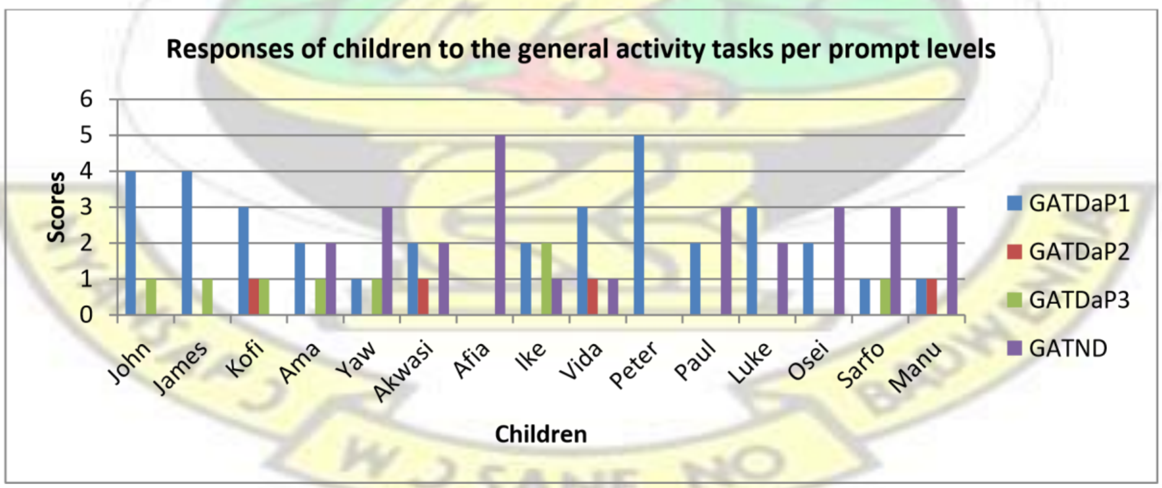


Figure 4.3 Responses of children to the general activity tasks per prompt

GATDP1: General activity tasks done at first prompt

GATDP2: General activity tasks done at second prompt
GATDP3: General activity tasks done at third prompt
GATND: General activity task not done

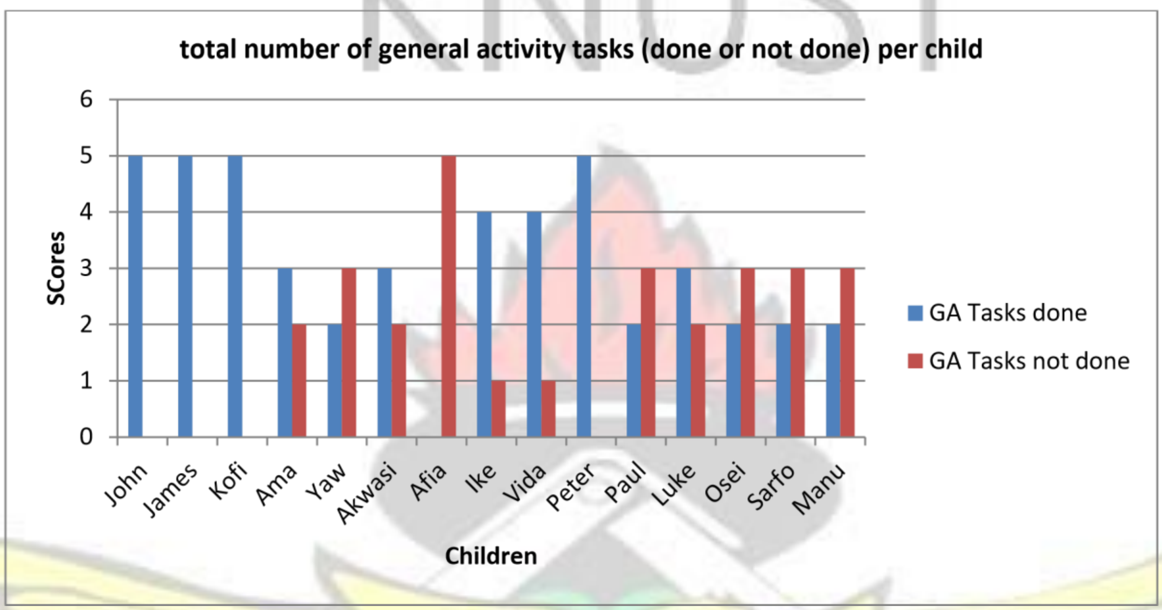


Figure 4.4 total number of general activity tasks (done and not done) per child

From figure 4.4, it is observed that 4 out of the 15 children performed all the 5 general activity tasks, 2 of them completed 4 tasks, 3 children did 3 tasks, 5 did 2 tasks and only 1 child did not do any of the general activity tasks. Compared to the im itation tasks, the number of children who did not do any of the general activity tasks is lesser (1 child) than the number that did not do any of the imitation tasks (5 children).

Table 4.2 Total number of responses for each general activity task

Command	Child responded when robot called him/her by name	How are you?	What is your name?	Take the ball	Give me the ball
Total number of correct responses for each GA task(out of 15)	13	12	8	8	6

From table 4.2, 13 of the children responded (by either turning to face the robot, walking towards the robot or saying "yes" or "hi") when robot called the child by name. The second most executed task was "how are you" to which there were 12 responses. The tasks "what is your name" and "take the ball" all had 8 responses and the task with the least response was "give me the ball". Perhaps due to the fact that their caregivers call the children by their names and also during their "Good morning" assembly sessions at school, each child is asked "how are you"? they were familiar with the first two tasks thereby accounting for the high responses. As some of the children are non-verbal, the decrease in responses to "what is your name" is not surprising. More so, the tasks involving taking the ball and giving it to the robot are unfamiliar and that could account for the low responses.

4.1.3 Physical contact with the robot

Autistic children easily concentrate and get stuck on objects of interest. During the interactions involving Rosye and the children, it was necessary to determine whether the children would like the physical appearance of the robot, get close to it and subsequently learn from it. Therefore, the experiments were carefully designed to begin with a quiet phase so as to observe the initial reaction of each child upon entry into the experimental room. Specifically, the effects of the uncanny valley phenomenon on the child-robot interactions were observed. Although Rosye is a humanoid robot, its features have been carefully designed to reduce its resemblance to human beings. From the video recordings, only 1 female child out of the 15 children expressed visible signs of fear upon seeing the robot for the first time. All the other children did not feel creeped out by the appearance of the robot; some smiled, touched and hugged it. This observation is backed by data in table 4.3 which presents the various parts of the robot touched by each child and the total number of times the children touched the robot.

In all, the robot was touched 193 times during the 15 child-robot interaction sessions. Whereas some of the children were more interested in engaging the robot in the imitation and general activities, others found delight in touching various parts of the body. Some children on the autism spectrum are sensitive to touch and therefore they may not like touching objects or also dislike being touched by objects or people. The high scores for "touch" indicated that the children were not intimidated by the look of the robot.

Table 4.3 Analysis of the number of times each child touched specific parts of Rosye

	Number of times child touched the robot's specific parts	Total

Child	Eyes	Neck	Mouth	Hand	Shoulder	Midsection	Head	number of times child touched Rosye
John	0	0	2	13	5	0	3	23
James	0	0	0	0	0	0	0	0
Kofi	0	0	0	2	2	0	2	6
Ama	0	0	0	8	1	6	3	18
Yaw	4	4	7	2	5	3	4	29
Akwasi	1	1	1	19	5	1	4	32
Afia	0	0	0	0	0	0	0	0

Ike	0	0	0	4	0	2	0	6
Vida	0	0	0	0	0	0	0	0

Peter	0	0	0	0	0	1	1	2
Paul	0	1	0	1	2	2	0	6
Luke	0	0	1	23	2	7	2	35
Osei	0	0	0	0	0	0	0	0
Sarfo	0	0	0	8	2	3	0	13
Manu	0	0	1	15	2	0	5	23
Total	5	6	12	95	26	25	24	193

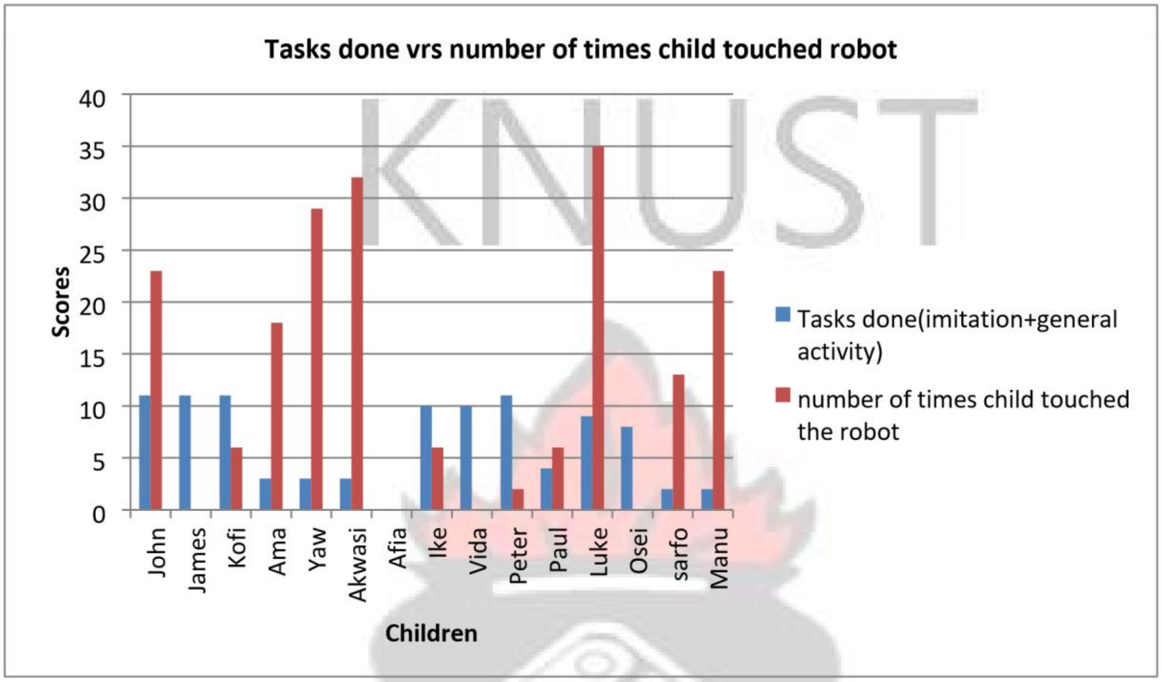


Figure 4.5 Number of tasks completed successfully by each child(maximum points achievable=11) vs. number of times every child touched the robot (child could touch the robot limitlessly)

From figure 4.5, some children who achieved lesser scores during the imitation and general activities touched the robot more times than those children who had higher scores during the games. 1 child, Afia neither touched the robot nor engaged in any activity; this was the child who was scared of the robot. In general, the preliminary study indicated that most of the participants were comfortable with the look of the robot. As such, the uncanny valley effect had minimal impact due to the fact that the autistic children touched Rosye multiple times during the interaction sessions and engaged it in numerous activities.

Table 4.4 Brief overview of results from the preliminary study

	Total number of tasks done per prompt			Total number of tasks done (all three prompts)	Total number of tasks not done
Tasks	1st prompt	2nd prompt	3rd prompt		
Imitation	40	11	0	51	39
General Activity	35	4	8	47	28

Table 4.4 shows that the total number of tasks done in the imitation and general activities are higher than the total number of tasks not done. 56.7% of the imitation tasks were successfully completed by the children whereas the percentage of imitation tasks not done was 43.3%. For the general activity tasks, 62.7% were successfully completed while 37.3% of the tasks were not done. The repetitive capabilities of the robot would make it a beneficial assistive technology for caregivers of children with autism who may get tired or frustrated repeating instructions to the children multiple times.

Results from the single day study indicated that the robot was able to engage some of the autistic children in imitation and general activity games and also succeeded in persuading some “uncooporating” children to perform the robot’s requests by giving several prompts. This preliminary experiment served as a baseline for investigating and assessing the responses and engagement levels of the children to the robot during the subsequent longitudinal studies.

4.2 Longitudinal study of interactions among Ghanaian autistic children and Rosye, a humanoid robot

The reaction of children on the autism spectrum towards play objects could be unpredictable; an object which was once their favourite could irk them on a different day. In the preliminary study, the attitudes and responses of the children towards an unfamiliar robot were assessed. Feedback from the children and their caregivers indicated that the robot appealed to and encouraged the children to participate in the imitation and general activity tasks. These positive reactions were quite surprising because according to the caregivers, most of these children seldom approach strangers (human beings) and get close to them like they did to the robot during an initial encounter.

The preliminary study (section 4.1) established a baseline (initial reaction of the children towards a novel humanoid robot). To investigate the effects of the "diminishing novelty effect", longitudinal studies were undertaken to assess whether their attitudes and reactions to the robot would be same, better or worse with time. Seven (7) autistic children, all of whom were from Autism Awareness Care and Training Center and had already partaken in the first experiment partook in experiment 2. Five (5) of them were males and the other two (2) were females. Six (6) of the children were verbal and one (1) was nonverbal. Five (5) of the children had the tendency to ignore their own names when called and all the children sometimes ignored instructions irrespective of who was issuing them. The mean age was 10.43 and standard deviation 1.62. The experiments were structured as individualized child-robot interaction sessions and the duration of each interaction session was flexible to accommodate each child's needs. The experimental script played by the robot was the same as the script used in experiment 1 as per the request of the caregivers.

A total of four hundred and two minutes twenty -six seconds (402m 26s) of video recordings containing interactions of the autistic children with the robot over eight sessions were analysed. Some of the children were absent at school on certain days and therefore missed a few experimental sessions. The results gathered from each autistic child-robot interaction was scored as follows: Each task (both imitation and general activity) had a score of 1 when it was done correctly by the child and a score of zero (0) if the child failed to correctly perform the task after three prompts. Quantitative and qualitative results of each child-robot interaction are presented as follows.

Child 1: James

James is a nine (9) year old male child with autis m. He follows instructions given by his caregiver but finds it difficult to approach, interact with and obey instructions from unfamiliar people. His responses to the imitation tasks were consistent except in the seventh session. Figure 4.6 depicts the pattern of the responses of James in both imitation and general activity tasks.

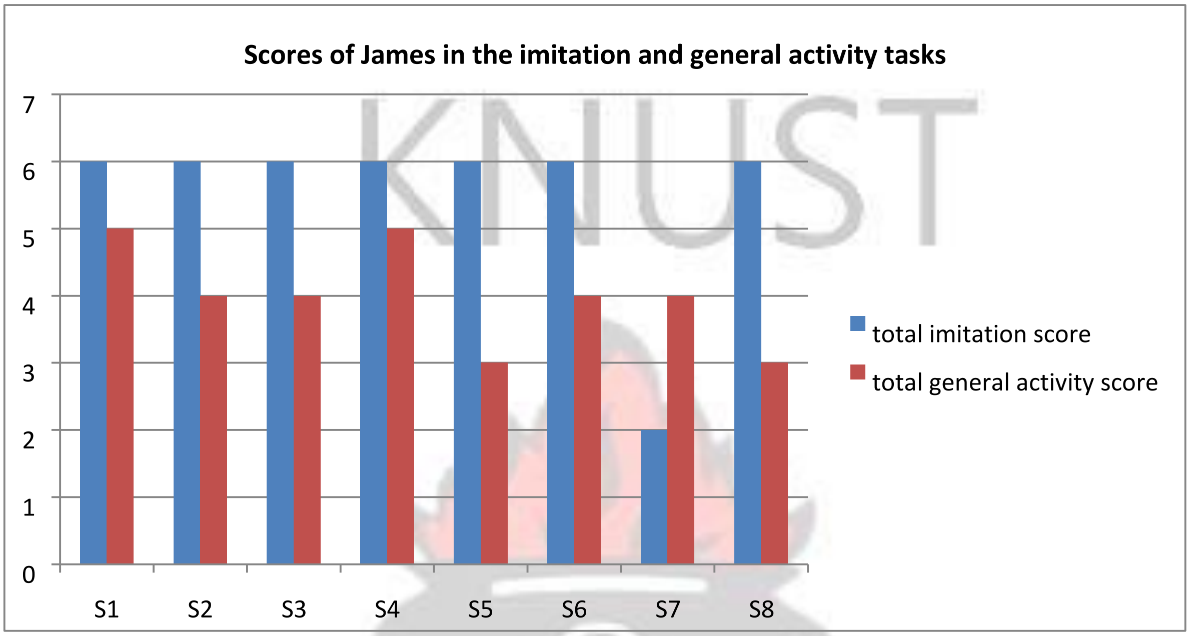


Figure 4.6 Pattern of the responses of James in both imitation and general activity tasks

James's scores for the imitation and general activity tasks were 44 out of a maximum of 48 and 32 out of a maximum score of 40 respectively. James touched the robot one hundred and forty-two (142) times, the highest recorded value compared to the number of times the other children had physical contact with the robot. After two interaction sessions, James became familiar with the routine of the games presented by the robot. For instance, he learnt the pattern of the actions in the imitation games and was able to follow the robot quickly in those tasks. Similarly, in the general activity games, he also learnt that after the robot asks him to take the ball, the next instruction would be for him to hand over the ball to the robot. As a result, in some of the sessions, he picked the ball when instructed by the robot and then handed it over to the robot (without being told to do so).

Child 2: John

John is a thirteen (13) year old male child with autism. He is verbal, likes to shout and yell for no reason and mostly ignores instructions, especially from strangers. He was present for seven out of the eight experimental sessions. In two of his sessions, he touched the robot, hugged it and called out “robot”. Figure 4.7 depicts the pattern of responses to the imitation and GA tasks for John.

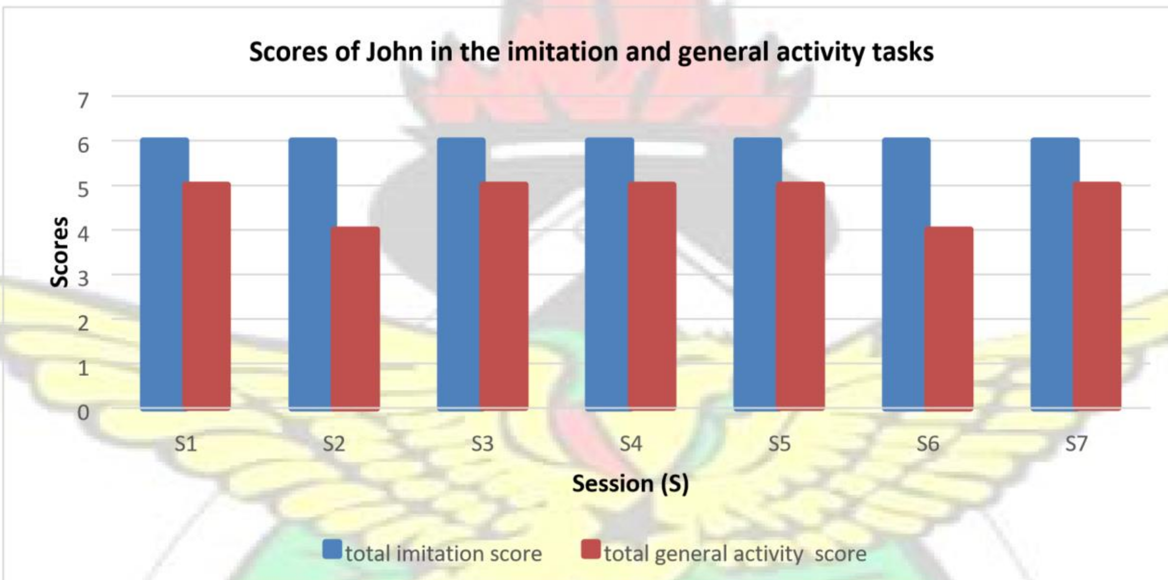


Figure 4.7 Pattern of the responses of John for both imitation and general activity tasks

John had one hundred and twenty-five (125) physical contacts with the robot, making him the second person who touched the robot the most. He scored 42 out of the maximum of 42 score for imitation tasks (for seven sessions) and 33 out of the maximum score of 35 for the general activity tasks (for seven sessions). John also became familiar with the robot’s actions, was always full of smiles and was quicker to engage in interactions with the robot.

Child 3: Ama

Ama is a thirteen-year-old verbal female with autism. She is calm, likes to smile and easily approaches strangers. She however sometimes fails to follow instructions. She partook in the experiment for seven sessions. She was able to reciprocate the robot's imitation actions from the fourth to the seventh sessions. Ama had a score of ten (10) out of the maximum of forty-two (42) for the imitation tasks over seven sessions. For the general activity tasks, she scored twenty-six (26) out of thirty-five (35) marks. Figure 4.8 depicts the pattern of responses of Ama to the imitation and general activity tasks over seven sessions.

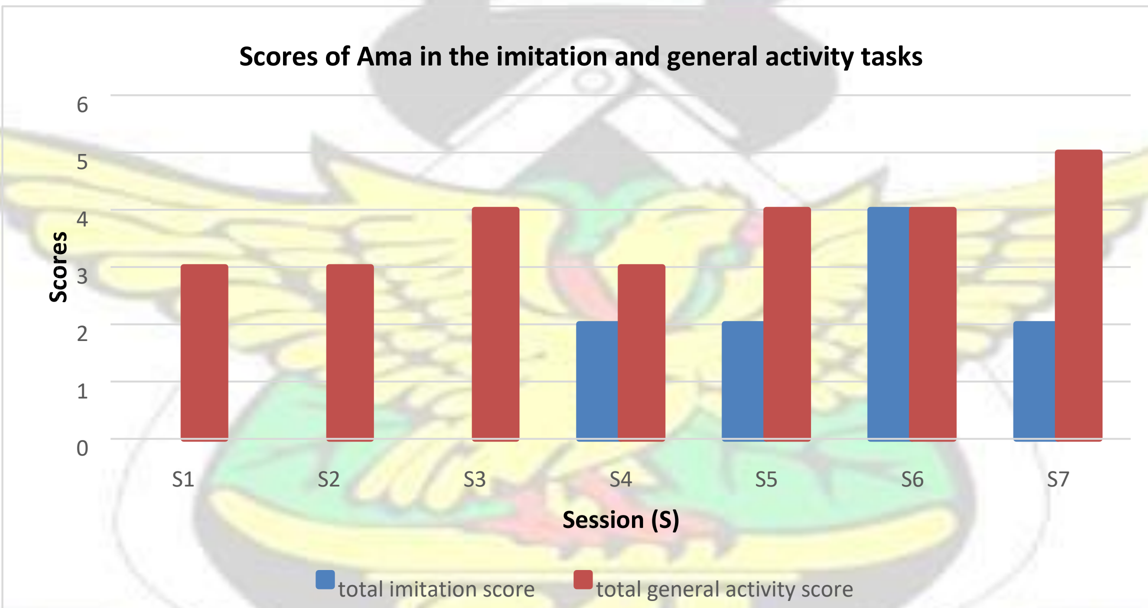


Figure 4.8 Scores of Ama in the imitation and general activity tasks

The frequency of her physical contact with the robot reduced as the sessions progressed over the days. Ama touched the robot forty-five times during her interactions over seven sessions. She sometimes repeated the robot's instructions instead of performing them; this trait is called echolalia, a common characteristic of somechildren on the autism spectrum. Ama also familiarized with the robot over time and from the third session onwards, whenever the caregiver informed her it was time to play with robot, she left her classroom for the experimental room without anyone directing her.

Child 4: Kofi

Kofi is a thirteen (13) year old child with autism. He is verbal and like other children on the spectrum, he sometimes ignores instructions. He consistently performed all the imitation and general activity tasks and touched various parts of the robot up to the fourth session after which there was a decline in his responses. In all the sessions, he would draw closer to the robot, touch its parts and return to a corner of the room to play with other toys. However, it was observed that even when his attention seemed to be off the robot, he was actually listening to it. Especially for the general activity tasks, he answered some of the questions (while he was still playing with an object) and for the tasks which required actions, he sometimes got up to perform the needed action. Figure 4.9 depicts the pattern of responses of Kofi to both imitation and general activity tasks.

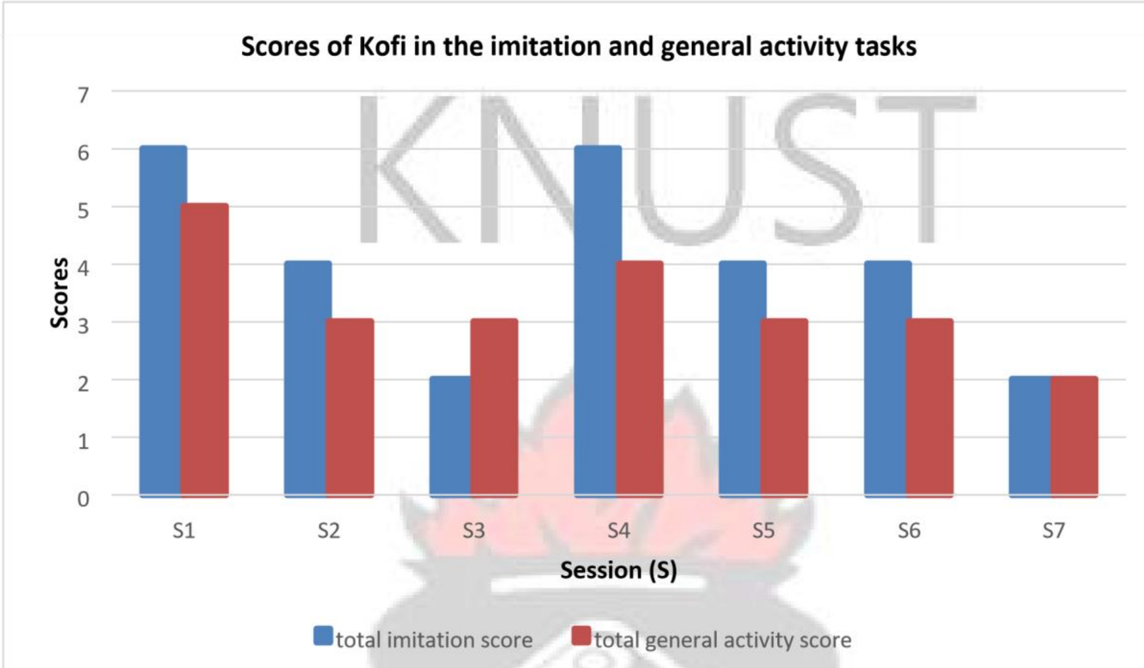


Figure 4.9 Scores of Kofi in the imitation and general activity tasks

His score for the imitation tasks was twenty-eight (28) of a maximum of forty-two (42) over seven sessions. For the general activity tasks, he obtained a score of twenty-three (23) out of a maximum of thirty-five. He touched the robot ninety-five (95) times during his interaction over seven sessions.

Child 5: Akwasi

Akwasi is a ten-year-old male child with autism. He is verbal, tends to ignore his name and instructions most of the time and has difficulty sustaining attention to learning tasks. His session with the robot lasted five sessions because he did not report to school on the other sessions. He always danced and sang along to the songs played by the robot. He did not respond correctly to any of the imitation tasks throughout his five-day interaction with the robot. His responses to the general activity tasks varied over the days. Figure 4.10

depicts the pattern of responses of Akwasi to the imitation and general activity tasks. He touched the robot thirty-seven (37) times in five sessions and spent forty-eight minutes twenty-eight seconds (48m 28s) interacting with the robot.

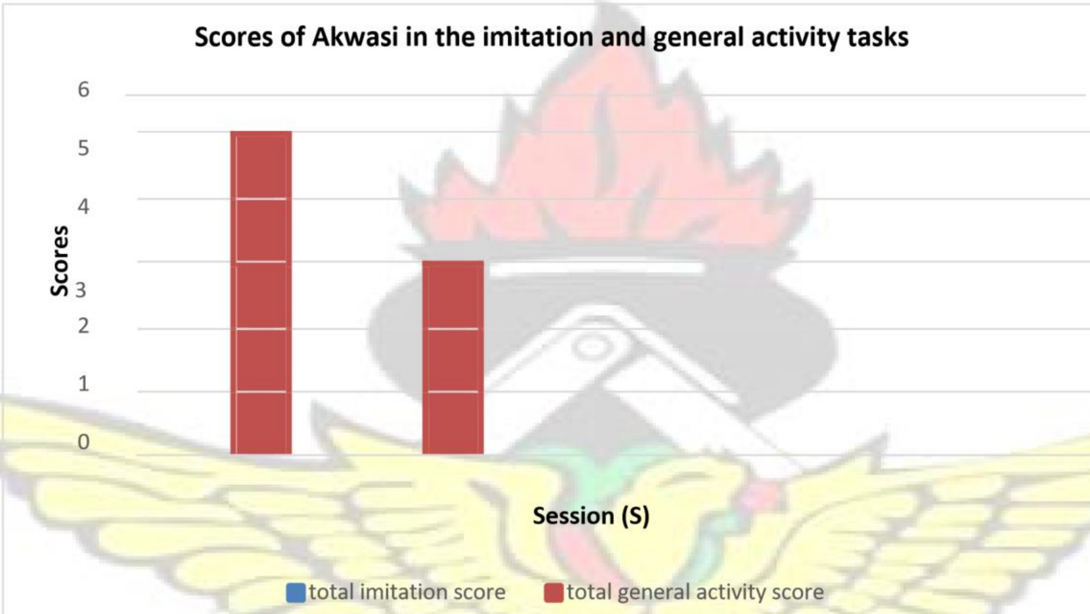


Figure 4.10 Scores of Akwasi in the imitation and general activity tasks

Child 6: Yaw

Yaw is a nine-year-old male child with autism. He is nonverbal, ignores his name, instructions, yells and hits others for no particular reason and has difficulty sustaining attention to tasks. He scored quite low in the imitation tasks. In most of the sessions, he turned his attention to the robot when it called his name. Notwithstanding the fact that he is nonverbal, he responded to "how are you" by using sign language. Figure 4.11 depicts the pattern of responses of Yaw to the imitation and general activity tasks. His overall score for touching the robot was sixty-eight (68).

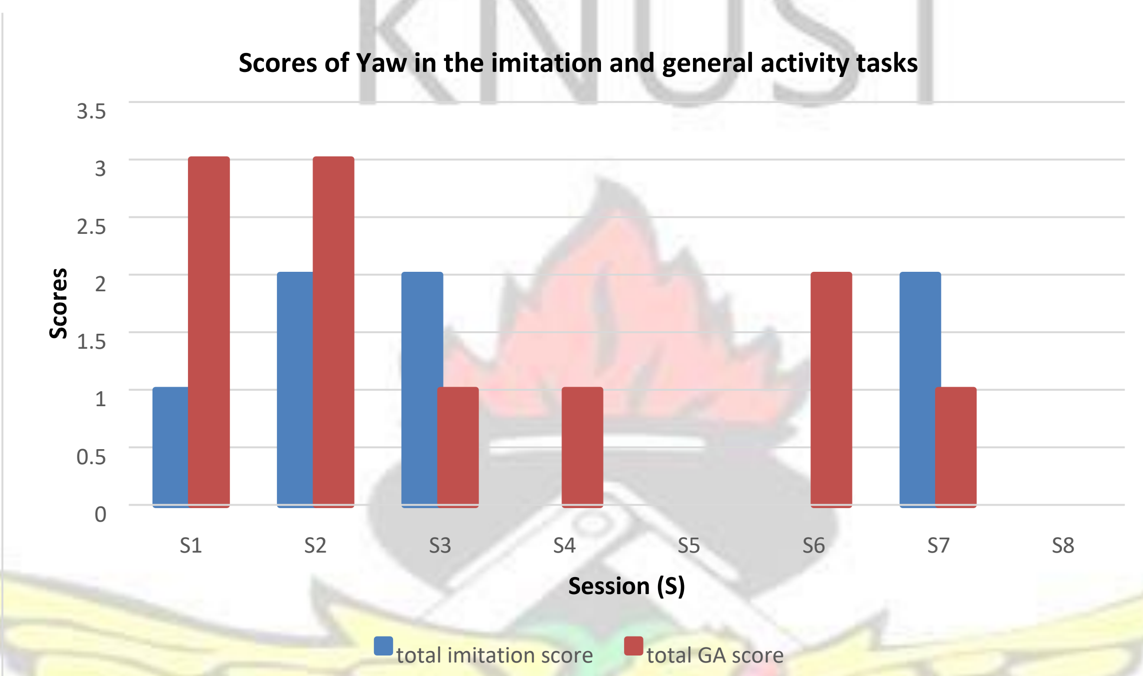


Figure 4.11 Scores of Yaw in the imitation and general activity tasks

Child 7: Afia

Afia is a nine-year-old female child with autism. She is verbal, sometimes ignores name, instructions, yells for no reason and has difficulty sustaining attention to tasks. She expressed visible signs of fear upon sighting the robot for the first time but surprisingly, she always wanted to come to the experimental room. She never drew close to the robot and got scared when the robot began the imitation tasks. From session 2 onwards, she stayed afar and danced to songs played by the robot. She was neither able to do any of the imitation nor the general activity tasks but spent forty-one minutes nine seconds dancing and watching the robot during the experimental period.

4.2.1 Physical contact with the robot

Figure 4.12 indicates the frequency of each of the seven children's physical contact with the robot.

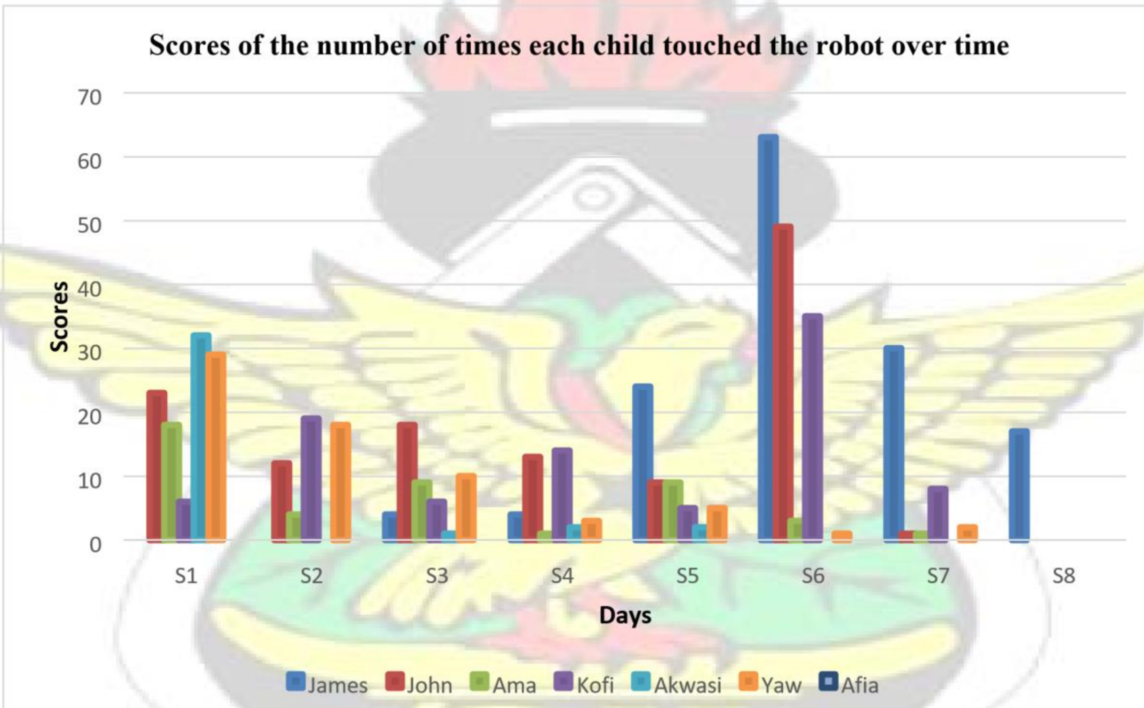


Figure 4.12 Scores of the number of times each child touched the robot over time

Figure 4.13 compares the responses of the children to the tasks as against the number of times each child touched the robot.

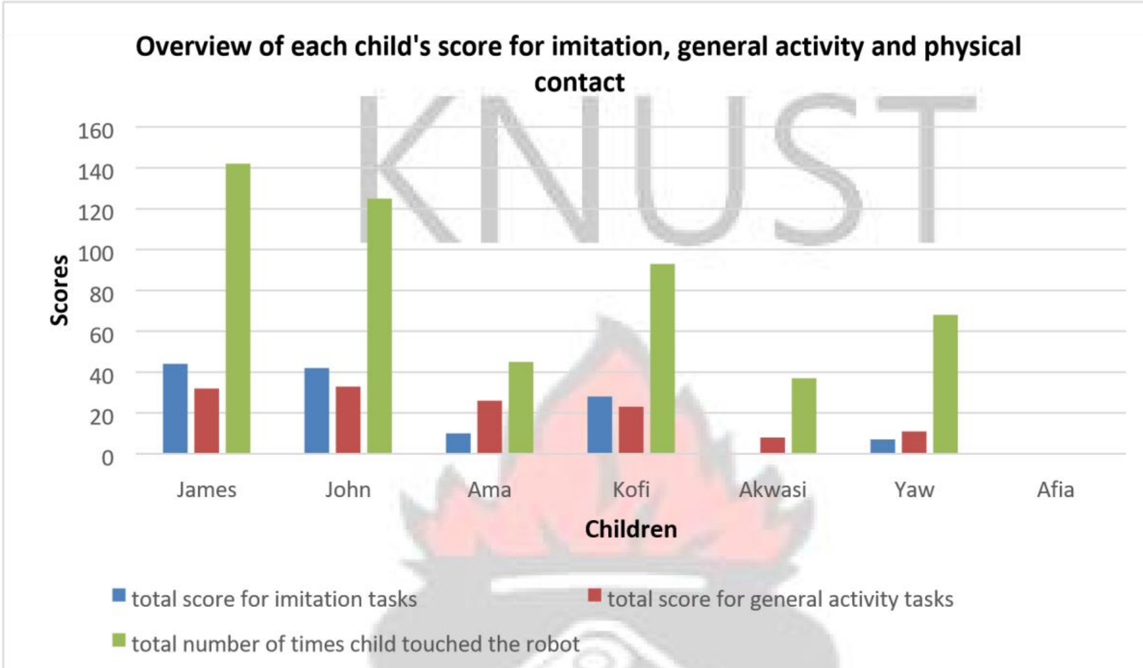


Figure 4.13 Comparison of the children's responses to both tasks (imitation general activity) and the number of times each child touched the robot

Throughout the experiments, the children touched various parts of the robot; eyes, mouth, head, midsection, shoulder, neck and hands. Some of the parts were touched more than others. The hands (both left and right) were the parts touched the most with a score of two hundred and forty (240). The second most touched part was the robot's midsection with a score of one hundred and four (104). The third was the head with a score of seventy-two (72), followed by neck and mouth with a score of twenty-nine (29) each. The shoulders of the robot were touched twenty-five times and the eyes were touched the least with a score of eleven (11). The robot was touched five hundred and ten (510) times during the experimental sessions. Over the eight session period, two hundred and eighty two (282) imitation tasks were presented by the robot to the children. The total score for the imitation tasks completed by the children was one hundred and thirty-one (131) whereas number of imitation tasks not done was one hundred and fifty-one (151), indicating that the number of tasks which were not done slightly outweighed the number of tasks completed. The robot presented two hundred and thirty-five (235) general activity tasks to the children. Out of this number, one hundred and thirty-five (135) tasks were successfully completed while the remaining hundred tasks were not done by the children. As opposed to the overall imitation results, the children performed better in the general activity tasks.

The patterns of responses of the children to the imitation, general activity tasks and the frequency of physical contact with the robot during the longitudinal studies have been presented in this section. Some of the children learnt from and became familiar with the robot's activities and therefore were able to request actions from the robot via speech or sign language. During the course of interactions, the children exhibited a few of the autistic traits they usually exhibit in their classrooms when being taught by the caregivers.

4.3 RoCA, a humanoid robotic assistive technology for children with autism

4.3.1 Features of RoCA

i. Communication and robot control: The robot has an integrated wifi and an embedded webserver which enables remote control via a computer or a mobile phone.

ii. Movement: The robot is able to raise both left and right hands up and down as well as turn neck left, right, up and down. Degrees of freedom (DoF): Head – 2, left arm -

1, right arm – 1.

iii. Speech functionalities

a. Voice output: The robot outputs sound with the aid of inbuilt speakers. Prerecorded sounds can be played by the robot. Output of sounds by the robot in real time is possible by speaking directly into the microphone of the computer or the mobile phone being used to control RoCA.

b. Speech processing

Through speech recognition functionality, the robot can be controlled by giving voice commands. Selected words or phrases picked up by the microphone can be analysed and the appropriate action(s) can be performed by the robot. RoCA can produce human voice through a speech synthesis module where words or phrases typed into the robot software control interface can be read out by the robot.

iv. Colour detection: The robot is equipped with a camera that has been configured to detect multiple colours.

v. Video streaming and live capturing of events: The robot's camera can record its environment and send live video streams over wifi to a remote computer.

vi. Modes of control

The robot can be controlled in two modes: Autonomous and Wizard-of-Oz (WoZ). In the WoZ mode, the care giver of the child can manually operate the robot and control the various modalities such as sound, movement and lighti ng. The caregiver can also select the various lessons to be delivered by the robot, repeat lessons or stop lessons when necessary. When put in autonomous mode, the robot can perform a pre-specified script which can include a combination of modalities such as movement, music and teaching lessons. The robot can also be instructed to sense the affective states of the children and decide on the appropriate activity to perform in order to sustain the interest of the children. In this mode, the robot always seeks the consent of the caregiver by means of a message prompt on the screen of the controlling interface, before changing lessons. Figures 4.14 and 4.15 indicate the front, back of RoCA and its motion capabilities.

Figure 4.14 Front and back view of RoCA

Both hands down

Left hand up

Right hand up

Figure 4.15 Motion capabilities of RoCA

4.3.2 Technical specifications of RoCA

The technical specifications of RoCA are listed in tables 4.5 and 4.6.

Table 4.5 Technical specifications of RoCA

Feature	Description
Dimensions	Height (mm): 1210 Width (mm): 480 Depth (mm): 425
CPU	220MHz 32-bit processing with two 32-bit ARM Cortex processors
Energy	Battery (LITHIUM-ION)
Interaction	Loudspeakers Microphones Wireless Cameras LEDs
Connectivity	Wi-Fi connectivity (ad-hoc/infrastructure/WEP/WPA/WPA2)
Sensors	Sonar

Table 4.6 Degrees of freedom of RoCA

Parts/Joints	Motor Type	Quantity	Degree of Freedom (DoF)	Coordinates
Head	HDD Standard Servo	2	2	Pitch (x-axis), Yaw (y-axis)
Left Arm	HDD Standard Servo	1	1	Yaw (y-axis)
Right Arm	HDD Standard Servo	1	1	Yaw (y-axis)
Base Wheels	Continuous Rotation Servo	2		360 degrees

ROCA is powered by Ez-Bv4/2 Wifi robot controller which has 24 digital/servo ports, 8 analog ports, 3 UART ports and 3 I²C ports. Figure 4.16 illustrates the Ez-Bv4/2 port layout.

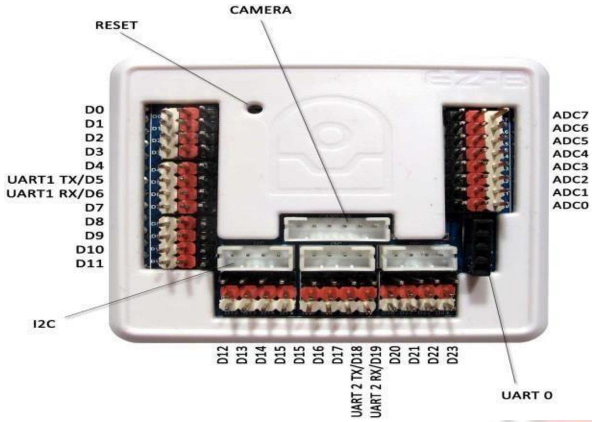


Figure 4.16 Ez-Bv4/2 port layout

4.3.3 RoCA's Software Control System

The robot was programmed using two IDEs: EzBuilder and Microsoft Visual Studio 2017 using the EZScript and C# programming languages respectively. The multi -tab custom interface for the robot which contains the motion, l essons, exercises, emotion detection and fuzzy inference modules. Figure 4.17 shows a screenshot of the custom robot control interface.

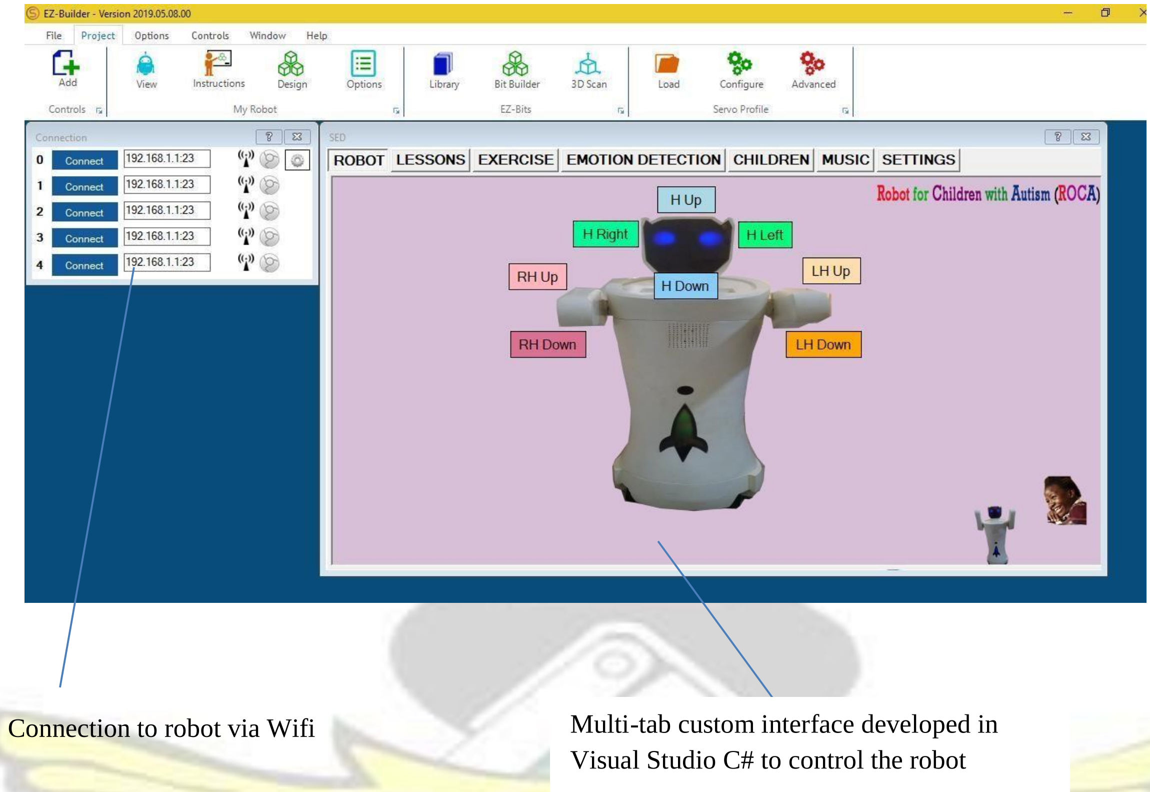


Figure 4.17 Custom RoCA robot control interface.

4.4 SingleShot Emotion Detector for Social Robots

In autistic-child robot interactions, the issue of fast real time and accurate analysis of the child's facial expressions picked by low resolution cameras onboard social robots still persist. Traditional emotion recognition systems separate face detection and classification pipelines making them less suitable for robots due to their exhaustive computational requirements, high latency, coupling and low frames per second(FPS). This thesis proposes a novel emotion recognition model, SingleShot emotion detector (SED) based on transfer learning for deployment on social robotic systems. In the proposed approach, object detection is treated as a classification problem where a single CNN is able to detect human faces, localize the face with bounding boxes and predict the emotional states of the detected faces. The proposed model combines face detection and classification into a single pipeline in order to provide high frames per second(FPS) and low latency for realtime emotion detection tasks in ACRI. The SED worked accurately on both high resolution and low resolution cameras. The developed emotion recognition model was first tested on a computer using its inbuilt web cam. Figure 4.18 depicts sample facial expressions taken from different angles by a PC webcam and their predicted classes given by the SED model. The localization is represented by bounding box coordinates and the classification is indicated by the predicted class of the emotional expression in the input image i.e. happy, sad, neutral, surprised, angry or fear. Figure 4.19 also shows some predictions by the model fed with images from the low resolution camera onboard RoCA.

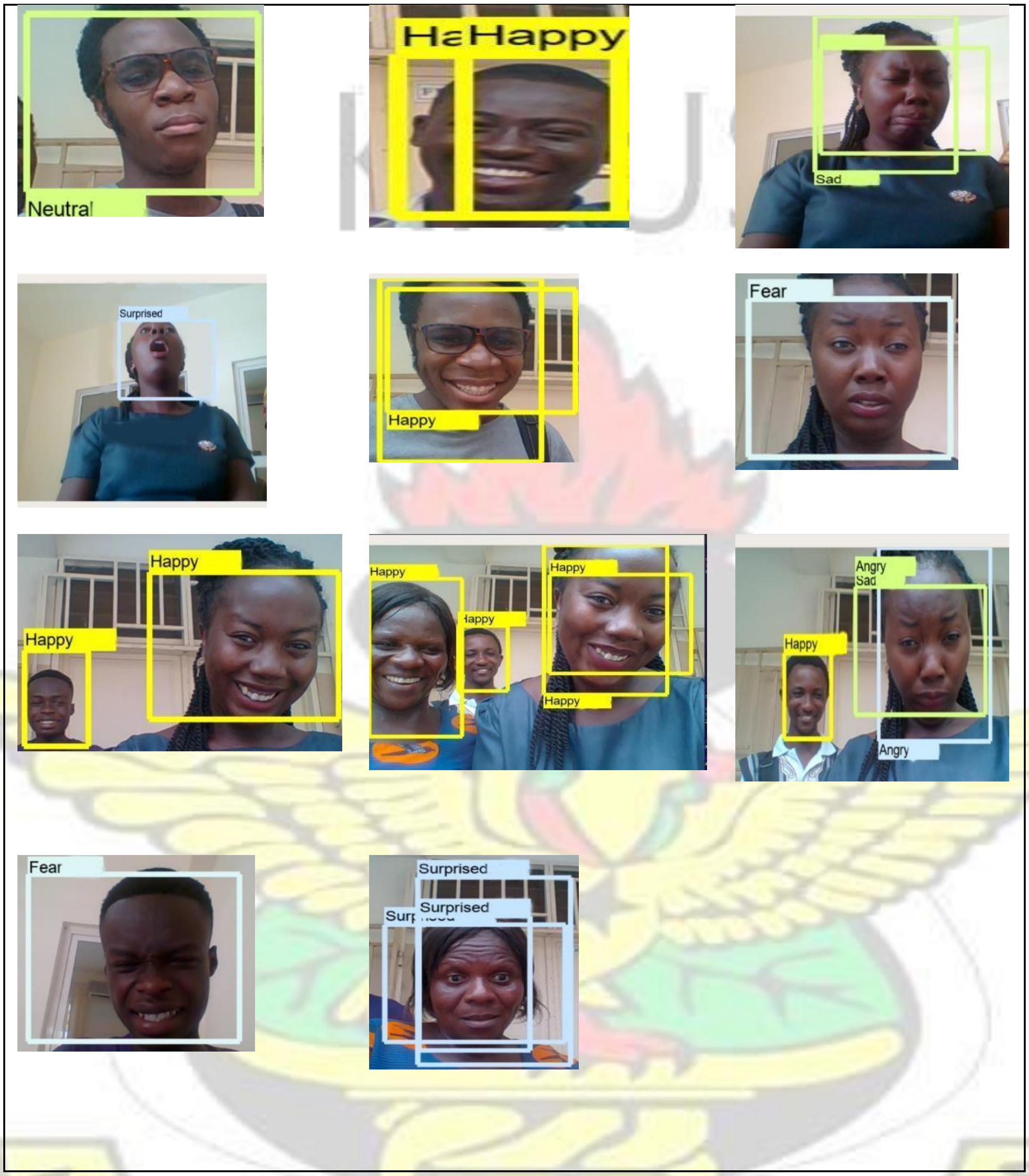


Figure 4.18 Sample facial expressions taken by a PC webcam and their predicted classes given by the SED model

After obtaining the desired ac curacy, the model was integrated into RoCA robot which has a low resolution camera and the second evaluation of the model was performed.

Figure 4.19 Sample images taken from the low-resolution camera onboard RoCA and their predicted classes

4.4.1 Evaluation metrics

Mean average precision (mAP) is a widely used technique to measure the accuracy of object detection models. mAP ranges from 0 to 1; a higher mAP indicates that the model is better at detecting multi classes. mAP of the SED model was automatically calculated by the Tensorflow object detection API and mAP of 0.93 was achieved (table 4.7). The total loss for the trained model was 0.16.

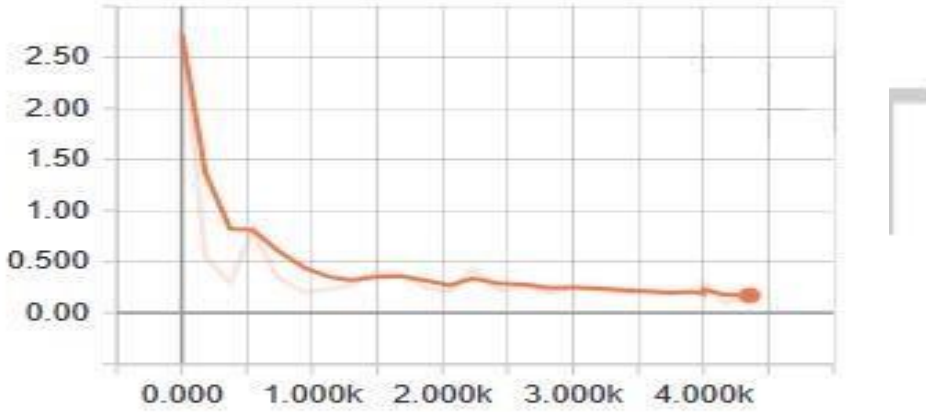


Figure 4.20 Total loss for the trained SED model

Table 4.7 mAP for the SingleShot emotion detector

Emotional Classes	Average Precision (AP) at 0.5IOU
Neutral	0.98
Happy	1.0
Sad	0.83
Fear	0.92
Surprised	1.0
Angry	0.85
Total AP	5.58
mAP (total AP of all emotional classes/6)	0.93

In table 4.7, the sum of the average precisions for all the emotional classes is 5.58. Hence the mean average precision of the Singleshot emotion detector is 0.93. The average precisions for the classes neutral, happy, sad, fear, surprised, angry are 0.98, 1.0, 0.83, 0.92, 1.0 and 0.85 respectively.

4.5 Effects of personalization on learning outcomes in robot mediated therapy for autistic children

The third experiment investigated the effects of personalization on learning outcomes in robot-mediated learning. The robot RoCA was programmed to w ork with two different setups, personalized robot and a non -personalized robot. When the robot operated in a non-personalized mode, the robot delivered the instructions to the children without considering factors such as aggregated emotional state, task difficulty and IQ level. In the personalized robot setup, the robot was able to consider factors such as the emotional state of the child (calculated by finding the most dominant emotional class logged to a file by the SED model) and task difficulty. The rob ot then used the available information to decide on the next appropriate action to take. The robot operating in the personalized mode ensured that each child’s learning sessions could be interspersed with appropriate tasks depending on prevailing situations.

The research question was: *Would the scores of the ASD children in a “fruit learning* *lesson" be higher in the RSEDFuzzy based personalized robot group than a group learning with a wizard of oz operated robot?*

Before the learning sessions, the inputs to the fuzzy logic controller were set in the fuzzy control interface (figure 3.8). Participants engaged in a total of 108 "fruits learning" lessons with RoCA, with 9 interaction sessions per child. Each lesson had four tasks to be completed by the children:

i. Take apple
ii. Give the apple to your teacher
iii. Take banana
iv. Give the banana to your teacher

In order to measure the effects of personalization on learning outcomes, the initial knowledge of the children on the fruits were assessed in the pre-test phase and the knowledge gained after the learning sessions were assessed using post-tests. In the pretest and post test lessons, each task had a score of 5 when performed correctly by the child. Statistical significance of learning gains between the RSEDFuzzy and RWOZ groups were determined by Mann-Whitney tests. Table 4.7 presents the pre-test and post-test scores of the children in the two groups.

From table 4.7, the difference between the aggregated pre-test scores for both groups is 20. A Mann Whitney test performed on the pre-test scores for both groups indicated that was no significant difference in the pre-test scores between the two groups, U=13.0, P=0.37. The difference in the aggregated post-test scores for both groups was 55 (table 4.7). A Mann Whitney test performed on the post-test scores for both groups indicated that there was significant difference in the post-test scores between the two groups, U=3.0, P=0.012. These results indicate that the children who received the personalized lessons from RoCA performed better on the post tests, thereby confirming the positive effects of personalization via the affect-aware and fuzzy-based engagement prediction model.

Table 4.8 Pre-test and post-test scores of the children in the RSEDFuzzy and RWOZ groups

Condition	Subject	Pre-test	Post-test
RSEDFuzzy	1	15	20
	2	0	10
	3	0	15
	4	10	20
	5	0	20
	6	15	10
Aggregated RSEDFuzzy group test scores		**40**	**95**
RWOZ	7	10	10
	8	0	10
	9	10	10
	10	0	5
	11	0	0

	12	0	5
Aggregated RWOZ test scores		**20**	**40**

4.6 Behavioural intention of special needs teachers to adopt robots as assistive technology in the classroom

The perceptions and behavioural intentions of the caregive rs of autistic children to adopt robots in classrooms were investigated through a survey.

4.6.1 Hypothesis

The hypotheses tested were as follows:

H1: Performance expectancy will positively influence teachers' behavioural intention to accept social robots in special needs education

H2: Effort expectancy will positively influence the behavioural intention of teachers to accept social robots in special needs education

H3: Social influence will positively affect the behavioral intention of teachers to ac cept robots in special needs education

H4: Facilitating conditions will positively affect the behavioral intention of teachers to accept robots in special needs education

4.6.2 Results

Data was collected from caregivers of some Ghanaian autistic children. The reliability of each construct was measured using Cronbach's alpha as shown in table 4.9. The profiles of the respondents are presented in table 4.10, descriptive statistics of data from the questionnaire and the UTAUT constructs are presented in tables 4.11 and 4.12 respectively.

Table 4.9 Reliability analysis using Cronbach's Alpha

Construct	Number of Items	Cronbach's alpha
Performance expectancy	4	0.959
Effort expectancy	4	0.888
Social influence	3	0.947
Facilitating conditions	3	0.786

Source: Field survey (2018)

Table 4.10 Demographic details of the study participants

Demographic variable	Description	Frequencies (n)	Percentage(%)
Gender	Male	21	42
	Female	29	58
Age group	18-30 years	20	40
	31-40 years	27	54
	40-50 years	3	6

Educational background	Up to Senior High School (SHS)	14	28
	Tertiary(1st Degree)	30	60
	Tertiary (Postgraduate)	6	12

Source: Field survey (2018)

Table 4.11 Descriptive statistics of data from the questionnaire

Variable	Description	Frequency (n)	Percentage (%)
Do you use technological devices when caring for children with autism?	No	20	40
	Yes	30	60
If yes, list the gadgets used	Computers	24	48
	Phones	6	12
Have you operated a robot before?	Yes	20	40
	No	30	60
I would prefer a robot for children with autism to look like	Human	34	68
	Animal	8	16
	Toy	8	16
I would prefer the length of the robot to be	Height of an average adult	9	18
	Half the height of an average adult	29	58
	Childlike	12	24
	Autonomous	9	18

What level of autonomy would you want the robot to have	Semi-autonomous	34	68
	Wizard-of-Oz	7	14

Source: Field survey (2018)

Table 4.12 Descriptive statistics for the UTAUT constructs n=50

	Construct	Mean	Standard Deviation
Performance Expectancy (PE)	PE 1	4.26	1.139
	PE 2	4.26	1.084
	PE 3	4.00	1.161
	PE 4	4.06	1.077
Effort Expectancy	EE 1	3.38	0.967
	EE 2	3.76	1.061
	EE 3	4.02	0.937
	EE 4	4.06	0.867
Social Influence	SI 1	3.80	1.355
	SI 2	3.92	1.383
	SI 3	4.02	1.450
Facilitating Conditions	FC 1	2.76	1.153
	FC 2	2.94	1.114

	FC 3	3.28	1.294
Behavioural intention	BI	4.28	1.457

Source: Field survey (2018)

In order to test the relationships between the UTAUT constructs and behavioural intention, multiple regression tests were performed in SPSS. Tables 4.13, 4.15, 4.17 and 4.19 indicate the ANOVA results for the UTAUT constructs in relation to behavioural intention. The model summaries for the UTAUT constructs are shown in tables 4.14, 4.16, 4.18, 4.20.

Table 4.13 ANOVA results for multiple regression analysis for Performance expectancy

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	30.845	4	7.711	4.738	.003[b]
	Residual	73.235	45	1.627		
	Total	104.080	49			

A. Dependent variable: I intend to use robots in special needs education in future B. Predictors: robots would make it easier for me to do my job, robots would be useful in special needs education, robots can increase productivity during school hours, robots would help accomplish teaching tasks more quickly

From table 4.13, performance expectancy significantly influences behavioural intention of special needs teachers to adopt robots in special needs education (p value of 0.03).

Table 4.14 Model summary, performance expectancy and behavioural intention

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.544[a]	.296	.234	1.276

The R square value of the model summary in table 4.14 indicates that performance expectancy contributed 30% to the total variation observed on behavioural intention of special needs teachers to use robots to teach autistic children in the classroom.

Table 4.15 ANOVA results for multiple regression analysis for Effort Expectancy

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	22.440	4	5.610	3.092	.025[b]
	Residual	81.640	45	1.814		
	Total	104.080	49			

A. Dependent variable: I intend to use robots in special needs education in future
B. Predictors: (constant), my teaching sessions using the robot would be clear and understandable, I would find a robot easy to operate, it would be easy for me to learn how to use a robot, using a robot in my work environment would be easy

From table 4.15, it is determined that effort expectancy significantly influences behavioural intention of special needs teachers to adopt robots in special needs education (pvalue of 0.025).

Table 4.16 Model summary, effort expectancy and behavioural intention

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.464[a]	.216	.146	1.347

The R square value of the model summary in table 4.16 indicates that effort expectancy contributed 22% to the total variation observed on behavioural intention of special needs teachers to use robots to teach autistic children in the classroom.

Table 4.17 ANOVA results for multiple regression analysis for Social influence

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	31.411	3	10.470	6.628	.001[b]
	Residual	72.669	46	1.580		
	Total	104.080	49			

A. Dependent variable: I intend to use robots in special needs education in future
B. Predictors: (constant), I would use robots if people who influence my behaviour think I should use robots, my working environment supports the use of robots, I would use robots if people who are important to me think I should use robots for teaching

As shown in table 4.17, social influence significantly influences behavioural intention of special needs teachers to adopt robots in special needs education (pvalue of 0.01).

Table 4.18 Model summary, social influence and behavioural intention

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.549[a]	.302	.256	1.257

The R square value of the model summary in table 4.18 indicates that social influence contributed 30% to the total variation observed on behavioural intention of special needs teachers to use robots to teach autistic children in the classroom.

Table 4.19 ANOVA results for multiple regression analysis for Facilitating conditions

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.702	3	.567	.255	.857[b]
	Residual	102.378	46	2.226		
	Total	104.080	49			

A. Dependent variable: I intend to use robots in special needs education in future
B. Predictors: (constant), a specific person should be available for assistance with robot operation difficulties, operation costs will encourage the use of robots, I have the required knowledge to operate a robot

As shown in table 4.19, facilitating conditions does not influence (pvalue of 0.857) behavioural intention of special needs teachers to adopt robots in special needs education.

Table 4.20 Model summary, facilitating conditions and behavioural intention

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.128[a]	.016	-.048	1.492

The R square value of the model summary in table 4.20 indicates that facilitating conditions contributed 16% to the total variation observed on behavioural intention of special needs teachers to use robots to teach autistic children in the classroom.

CHAPTER 5
DISCUSSION

This chapter focuses on addressing the research questions from the various experiments and discussions on the implications of this research.

5.1 Preliminary observations from interactions among Ghanaian autistic children and Rosye, a humanoid robotic assistive technology

Whereas typically developing children start to develop theory of the mind (ToM) from age 4 (Kloo et al., 2010), autistic children may or may never develop ToM (Barron-Cohen, 2000). Children with autism may not understand other people's emotions and intentions due to deficits in ToM. The inability of autistic children to develop ToM affects their social interactions with humans and other entities. ToM can also be linked to how people experience the uncanny valley phenomenon. A research question that needed to be addressed was: "*How would the Ghanaian children on the autism spectrum experience the effects of the uncanny valley phenomenon during their initial encounter with a humanoid robot Rosye?*

From the findings of the empirical study (section 4.1.3) indicate that out of the 15 children, only 1 female child expressed signs of fear upon seeing the robot for the first time. All the other children were not creeped out by the robot; some smiled, touched and hugged it. Whereas some of the children were more interested in engaging the robot in the imitation and general activity games, others found delight in touching various parts of the body multiple times.

Based on observations and findings from the preliminary experiment, the effects of the uncanny valley during the Ghanaian autistic children and robot interaction was minimal. In the case of Rosye, although it is humanoid, its features have been crafted in order for it not to have striking resemblance to humans and that could have contributed to the results obtained. This finding is in line with research by Ueyama (2015) which suggests children with autism do not resonate the uncanny valley effect.

Imitation plays a vital role in cognitive development and social communication. Typically developing children are able to acquire new skills through imitation of adults and young ones. Children with autism exhibit imitation skills deficits and as such, they find it difficult to socially communicate, learn and interact with others. If an autistic child is taught imitation skills, there is likely to be improvement in his or her cognition and social development. A current challenge lies in identifying suitable strategies to teach imitation to these children. This thesis probed the effectiveness of robot-mediated imitation practice for Ghanaian children with autism. One of the goals was to test whether a humanoid robot can engage the children in imitation games.

The results indicate that more than half of the participants achieved a 100% imitation response score, indicating they successfully imitated all the six actions performed by the robot. One-third of the children did not imitate any of the robot's actions and a few imitated some of the robot's behaviour. Although mixed results were obtained, these results suggest that some Ghanaian autistic children are likely to respond to robotmediated imitation practice. These results are very encouraging and provide new level of "hope" for Ghanaian caregivers who many at times struggle to catch the attention of autistic children during imitation sessions.

Some children on the autism spectrum exhibit unusual reactions to strange objects and may resist changes in their routines. Others may experience sensory sensitivities to light, sound, touch, smell and taste. The interaction scenarios for the general activity tasks were carefully designed to investigate whether the children would listen to and obey instructions from the robot. A deliberate decision was made to individualize each interaction session by enabling the robot to refer to each child by name. From table 4.4, it is observed that 13 out of the 15 children responded (either verbal, walking towards the robot or turning to look at the robot) when the robot called each child by name, making it obvious that the children acknowledged the presence of the robot.

Another notable observation was the fact that, two of the non-verbal children responded to "how are you" using sign language. More than half of the children also mentioned their names to the robot. Among the general activity tasks was one which was meant to elicit joint attention behaviours from the children. This was GA task 3, "*name of child* + take the ball". The robot pointed to a ball lying at a corner of the experimental room and it was expected of the children to look in the same direction (i.e. share a common focal point), and then go and pick up the ball. 8 out of the 15 children successfully completed this task indicating the success of the robot in eliciting joint attention behaviours from majority of the children. Overall, a higher cumulative score was recorded for the general activity tasks as compared to the imitation game. Rosye was able to serve as a communicative partner to these children, hence it could potentially server as a tool to aid in teaching social interaction and joint attention skills to autistic children.

Whereas some children on the autism spectrum may learn new skills easily, others often require multiple trial sessions and prompts. The non-intrusive verbal prompting strategy adopted by the robot proved to be effective. It is observed from figures 4.1 and 4.3 that some children were able to respond correctly to tasks after multiple prompts. This finding is in line with research by Bishop et al. (2019) which suggests that prompting is an effective intervention for children with ASD irrespective of their cognitive levels and communicative abilities. Humans have the tendency to easily experience mood swings and frustrations when they are made to repeat stuff multiple times. In contrast, a robot can behave in the same way limitlessly without changing emotions or getting tired. For children with ASD who may require several prompts in order to perform a task or learn a skill, social robots could be suitable aids for caregivers. The repetitive capabilities of the robot make it a beneficial assistive technology for caregivers of children with autism who may get tired or frustrated repeating instructions to the children multiple times.

5.2 Longitudinal study of interactions among Ghanaian autistic children and Rosye, a humanoid robot

The patterns of responses of the children to the imitation, GA tasks and the frequency of physical contact each child had with the robot in the longitudinal empirical study have been presented in sections 4.2. Consolidated data on the performance of the children in the imitation game indicates that the over an extended period, the number of tasks which were not done were slightly more than the number of tasks successfully completed by the children. On the other hand, for the GA tasks, the number of them completed successfully outweighed the number of tasks which were not done.

For many children involved in this experiment, continuous exposure to the robot had a positive impact on them. Three of them who scored high during their interaction with Rosye in the first session exhibited similar levels of engagement and enthusiasm over the subsequent days. One child was able to engage in the imitation tasks from the fourth session onwards. Another child did not response to the imitation task after the third session but responded to a few in the last session. Only one child showed disinterest and would not engage the robot over the sessions. The child who was initially scared of the robot overcame her fears and improved by smiling and dancing to songs played by the robot on subsequent sessions. These findings are in line with research by Scassellati et al. (2012), which indicates that the frequency of autistic child-robot interactions can have an impact on learning outcomes. From the longitudinal study undertaken, it can be observed that the children benefitted from extended interaction periods and confirms similar research by Robins et al. (2004), Kozima et al. (2007), Duquette et al. (2008), Wainer et al. (2010) and Valentina (2017).

The children also touched various parts of the robot multiple times during their interaction sessions and affection (hugging, smiling) was shown by a few of the children to the robot. Some of the children learnt from and became familiar with the robot's activities and therefore were able to request actions from the robot via speech or sign language. These results have indicated that children on the autism spectrum can through repeated interactions with social robots. During the course of interactions, the children exhibited few of the autistic traits they usually express in their classrooms when being taught by the caregivers.

To effectively function as an assistive technology to the children, a social robot should among other things be able to entertain, engage and more importantly teach the children new skills. Therefore, the research also sought to find out whether apart from responding to instructions posed by the robot, the children would be able to pick up new skills and also learn from the robot. Presenting the robot to the children over the extended interaction period and repeating the same tasks in this timeframe afforded the children the opportunity to become familiar with the robot. From the third session onwards, the pattern of hand movements made by the robot were picked by some children and they could follow through the imitation exercises easily.

For some of the children, they realized that the next task after "take the ball" was to give the ball to the robot. As a result, whenever these children were asked by the robot to take the ball, they would take it and hand it over to the robot even before the robot instructed them to do so. A few took notice of the experimental room and would come there themselves as soon as their caregivers told them it was time to play with robot. These indications have provided some evidence that some of the children learnt from the robot as a result of continuously engaging with it.

5.3 Prediction of emotions from facial expressions of autistic children

In human-robot interaction, knowing and understanding human's affective state is crucial for sustained long term interactions. Although facial emotion recognition, FER, has improved over the years, there are still setbacks hindering their successful deployment on social robots. Significant weaknesses in existing FER systems include low accuracy levels, lighting conditions and lack of datasets for relevant application domains. A lot of these FER systems need high end and costly GPUs for real-time deployment making them unsuitable for robotic applications and embedded systems.

Deploying deep learning FER algorithms for autistic child-robot interactions is a good idea. However, the problem lies in the unavailability of relevant datasets especially for protected groups like autistic children. The proposed system takes a new approach towards predicting emotional expressions of autistic children. The proposed SED model presents a new approach to FER for robotic systems by applying object detection to classify emotions of autistic children according to six emotional classes. This approach, addresses the issue of data deficiency, more accurate and fast real time predictions for use in robot enhanced therapy for children with autism.

In the proposed model, transfer learning has been performed on SSDLite to derive a new model for the target domain due to the sparsity of datasets. Knowledge in the form of weights and feature maps previously learnt in SSDLite were used to train the SED model on very small datasets to prevent over fitting and achieve the desired results in a shorter period of time. Since the SED model is based on SSDLite which uses depthwise separable convolutions and mobileNetv2 as feature extractor, it is optimized to ensure successful integration on social robots. The mAP achieved for the six emotional classes was 93%, a better accuracy than Liu et al.(2008) and Krupa et al. (2016) who obtained 82.9% and 90% accuracy respectively using physiological signals to classify emotions of autistic children with Support Vector Machines. The proposed model also outperforms an emotion recognition system for autistic children by Smitha and Vinod (2015) with an accuracy of

82.3% using images with Principal component analysis.

5.4 Effects of personalization on learning outcomes in robot mediated autism therapy

Experiment 3 aimed at empirically testing the effects of a personalized affect-aware and fuzzy-based engagement prediction model on learning gains in autistic child-robot interactions. The research question was: *Will the scores of the ASD children in a "fruit learning lesson" be higher in the RSEDFuzzy based personalized robot group than a group learning with a wizard of oz operated robot?*

The absence of significant difference in the pre-test scores for the RSEDFuzzy and RWOZ groups indicated that all the children began experiment 3 with similar knowledge on the two fruits, apple and banana. The post test scores for both groups confirmed an increase in learning gains. Some children who previously could not identify any of the fruits were able to do so while other children who scored low in the pre-test improved their performance in the post test. These results further support the results from experiment two of this thesis and research conducted by Duquette et al. (2008), Wainer et al. (2014) and Valentina (2017) which suggest that autistic children would benefit from longitudinal repeated lessons with robots. From table 4.7, although both groups recorded an increase in learning gains, a Mann Whitney test confirmed that the learning gains were higher in the RSEDFuzzy group which received personalized interactions from the robot. These results indicate that, creating rapport and sustaining engagement is crucial in autistic child – robot interactions. The affect aware and fuzzy based approach adopted by the robot led to the co-operation, prolonged bi-directional communication and increased engagement in the RSEDFuzzy group.

The main contribution to knowledge in this thesis is the proposed personalization framework to be deployed on and used by robots for autistic child – robot interaction sessions. This framework consists of a novel SingleShot emotion detector and a fuzzybased engagement estimation system programmed in Python and C# respectively. The framework has been evaluated on a humanoid robot RoCA by undertaking numerous autistic child – robot interaction sessions and results have confirmed that, when lessons are personalized, autistic children are likely to learn at a faster rate. Many other works on robot – mediated learning for autistic children typically used a single day session (Valadão et al. (2016); Kim et al. (2013); Shamsuddin et al. (2012); Robins et al. (2009); Stanton et al (2008)) or wizard-of-oz based longitudinal studies (Valentina (2017); Wainer et al. (2010); Duquette et al. (2008); Kozima et al. (2007); Robins et al. (2004)). More so, data from this research can further help researchers to conduct more studies on appropriate robot-mediated interventions for autistic children in resource constrained environments.

5.5 RoCA, a humanoid robotic assistive technology for children with autism

In the participatory robot design process, the stakeholders were actively involved in the requirement elicitation, design, identification of interaction scenarios and evaluation of the robot. This approach sought to contribute towards addressing a gap in autism research where social robots have been designed and developed for the children instead of being co-designed with them (Aslam et al., 2019). Unlike other social robots like Nao, Kaspar and Probo which were designed before being used for autism therapy, the participatory design approach adopted in this research has provided several advantages.

The contributions of the teachers gathered from the interviews and group discussions were influential in shaping the robot development process. For a domain like autism management, the choice of hardware and software for technological systems is very crucial due to ethical reasons and the fact that some of the children are likely to exhibit very aggressive behaviours. For the prototype of RoCA robot, the main body was carved out of polystyrene foam because it is low cost and less fragile. Significant differences can also be observed between Rosye and RoCA robots. Rosye is an adult humanoid robot whose outer coating consists of multiple colours. Through the various discussions with stakeholders, a decision was made to reduce the height of the robot to half the height of an average adult and also give it a neutral colour i.e. white. Children with autism more often learn through play and as such a smaller robot may likely to appeal more to them as compared to an adult which some children could find intimidating. However, this assertion also needs to be investigated further in future studies. From the first and second Wizard of Oz experiments, it was observed that only one child felt intimidated by the robot while all the others were comfortable around it, although the robot was of adult size.

The results of these experiments bring to light the fact that, one approach may not work for all the children. From the research, it has been identified that for robots to be successfully deployed and serve their purpose as assistive technologies in special needs education, three main factors have to be considered: scaffolding, adequate training for the teachers who would use the robot and associated costs. The results from this study reiterates the need for robots to be scaffolded and given different appearances and voices (mechanical, childlike or adult) based on the preferences of the child. These findings are in line with research by Robins et. al (2005) who opined that scaffolding is necessary in human-robot interaction. For instance, during a discussion with one occupational therapist who works with autistic children, she said "*there is this particular autistic child I know. He would not do any work unless you bring a particular picture of an animal close to him*". Such a child may respond well to the robot when pictures of the animal are wrapped around the robot or frequently displayed on the robot screen.

5.5 Behavioural intentions of special needs teachers to adopt robots as assistive technologies in special needs education

Apart from scaffolding, the one other major factor on which the successful deployment of robots depend is the behavioural intention of the teachers to adopt the robots in the classroom. The results from the behavioural intention survey conducted (section 4.6) depict the state of robot technology adoption in special needs autism education in Ghana. More than half of the survey participants indicated that they did not use technological devices in the classroom; a few of the teachers indicated they use phones and computers in the classrooms. None of the teachers had ever used a robot in the classroom for a teaching session. These results suggest that, Ghana is behind with regards to technology adoption in special needs education for autism management as compared to the Western world. Despite possible benefits associated with robot mediated therapy for autism, ethical issues which could arise include undesired attachment to robots, limited human contact, loss of privacy and loss of jobs.

In order to devise strategies for overcoming these concerns, the survey respondents were asked which level of autonomy they wanted social robots to have. From the responses, 68% of them preferred a semi-autonomous robot, 18% preferred an autonomous robot and 14% preferred a WoZ controlled robot. The low responses derived for the WoZ are not surprising because operating a robot in this mode indirectly means the teacher's attention would most of the time be on the robot instead of the child; and this would be an undesirable situation. The surveyed responses highlight that the teachers prefer the robots to have supervised autonomy where the teachers would have control over the robot and can intervene when necessary. The UTAUT model was also empirically tested using data collected from the teachers of autistic children to investigate the factors influencing their behavioural intention to adopt robots in special needs education. The findings indicated that performance expectancy (PE), effort expectancy (EE) and social influence (SI) significantly influences the behavioural intention of teachers to adopt robots in special needs education.

The responses from the survey (section 4.6.2) revealed the teachers agree that robots would be useful in special needs education by increasing productivity, efficiency and effectiveness. It was found that PE positively influences behavioural intention with p=0.03, F=4.738 and R square=0.296. Therefore, teachers who have high PE are likely to use robots in special needs education. This study is in line with research by Wong (2015) and Zalah (2018) who also have presented that teachers feel technology can enhance teaching and learning activities.

According to UTAUT, people's willingness to use technology depends on effort expectancy, that is how easy it would be for them to use it. The survey results for effort expectancy p=0.025, R square=0.216 and F=3.092 indicates that effort expectancy significantly influences behavioural intention. The teachers think it would be easy for them to learn to use, operate and deliver teaching sessions with the robot although more than half of the respondents indicated they had not used a robot before and do not have computer science or IT background. This is a beckoning call to human-computer interaction experts to adopt participatory design in the development of technologies for special needs education in order to ensure long term acceptability and ease-of-use. These results are consistent with research by Nair & Das (2012) which state that perceived ease of use influences behavioural intention to use information technology in teaching.

According to the results, social influence with p=0.001, F=6.628 and R square=0.302 significantly influences behavioural intention. This finding is in line with research by Radovan & Kristl (2017) who found positive relationships between social influence and behavioural intention. This means that the decision of the teachers to use robots may be influenced by the people around them. Facilitating conditions focused on the roles that external factors such as availability of training experts to assist the teachers with the robot usage and the operation costs of robots in Ghana would play in the willingness to adopt robots. Autistic children in reso urce constrained environments are likely to benefit from assistive technologies but more often, the associated costs with their purchase continue to hinder their deployment. The survey results p=0.857, R square=0.016 and F=0.255 indicates that currently, facilitating conditions in Ghana do not support usage of robots in special needs education. The teachers were of the view that they do not have the required knowledge to operate a robot. More so, the costs involved in robot mediated therapy may hamper the e ffective use of robots. However, they were willing to learn and operate the robots in the classroom if the necessary infrastructure is provided. These results suggest that, for robots to be successfully deployed in Ghana, the teachers have to be trained tobe able to use them and also the costs associated with the deployment and maintenance should be on a low side.

CHAPTER 6

CONCLUSION, RECOMMENDATION AND FUTURE WORK

Children on the autism spectrum are special groups of people who have unique learning needs and require extra support and care. Currently, there is mounting evidence on the suitability of technology-based interventions for such children. However, research has proven that adaptability, cultural backgrounds, costs and ease of use play a crucial role in the efficiency, acceptance and effective use of technology. Therefore, researchers ought to focus on approaches which are aimed at investigating the capabilities, challenges and needs of people with special needs who are often under-represented in the design of assistive technologies.

This thesis adopted a participatory design approach to investigate the suitability of a humanoid robot as an assistive technology for Ghanaian autistic children and propose strategies for personalization of robot-mediated learning sessions using deep learning and fuzzy logic. Numerous child-robot interaction experiments have been conducted over several months with children living with autism, having varying levels of cognitive, verbal and social dis(abilities). These experiments aimed at evaluating the impact of social robots and the effects of personalization strategies on the improvement of communication, social skills and joint attention skills of autistic children. A humanoid robot Rosye was used in the first two experiments: to investigate the initial reaction of some Ghanaian autistic children to an "unfamiliar" robot and secondly to assess the reaction and engagement levels of the children in interactions with the robot over an extended period of time.

Feedback from the first two experiments, information gathered from numerous visits to some autism centers and inputs from health care professionals and psychologists guided the development of the novel robot RoCA and the specification of interaction scenarios for the robot-mediated learning sessions. RoCA is equipped with wifi, ultrasonic sensor, speakers and a camera. It has motor and verbal skills, emotion detection, speech synthesis and speech recognition modules and a desktop interface for controlling the robot.

This thesis also addressed the issue of fast and accurate emotion prediction for autistic child-robot interaction. A transfer learning approach to emotion detection based on object detection has been introduced to address the issue of the sparsity of datasets in autism domain. This approach of using object detection in emotion classification is computationally efficient even with small datasets and also achieves fast frames per second and low latency.

RoCA was equipped with the proposed SingleShot Emotion Detector and a fuzzy-based engagement prediction model to semi-autonomously adapt the child-robot interactions to suit each child's needs. The personalized RoCA system used the SED model to detect the affective states of the autistic children and the fuzzy-inference to predict the estimated engagement levels of the children in order to adapt the learning sessions. A Mann-Whitney test performed to assess statistical significance of learning gains associated with personalized robot-mediated learning confirmed that the SED model and the fuzzyinference system collaboratively enhanced learning gains more than a wizard-of-oz operated robot.

The main contributions of the thesis are as follows:

i. development of a framework consisting of a novel SingleShot emotion detector and fuzzy-based engagement prediction inference system for personalized robotassisted learning for autism management ii. assessment of the reaction and engagement levels of some Ghanaian autistic children over extended interaction periods with a humanoid robot iii. assessment of the effects of personalization on learning outcomes in robotmediated learning iv. design and development of a low cost humanoid robotic assistive technology, RoCA for autistic children

The humanoid appearance and design of the robot seemed to have contributed to the enthusiasm expressed by the children since they saw it as “similar” to them. The robots were also able to persuade some “non-cooperating” children to respond to instructions via multiple prompted cueing. Presenting the robot to the children continuously enabled them to familiarize and learn from it as the sessions went by. Findings from this thesis suggest that most of the children have engaged and responded well to the robots. However, due to variability in autism manifestation among individuals, there is the need for robot and software customization to cater for individual preferences. Children with autism learn through repetition and consistency and robots are better situated to deliver the same tasks over and over again without getting tired or bored. As a result, robots can be promising tools to supplement the efforts of caregivers of autistic children. Robots for use in autism therapy need to be cost effective and easy to use by professionals and caregivers of these children with minimal IT background. With this cost effective robot which has been developed, more studies can be conducted to investigate diverse ways the robot can assist the children with other academic, life skills and sensory activities. Results from this thesis serves as a contribution to knowledge in research on robots in autism therapy, majority of which have been done in the developed countries.

6.1 Future work

The findings of this research are restricted to classroom based interactions. In future, more studies could be conducted to investigate the efficacy of social robots for the autistic children in clinical settings using standard therapy practices such as Applied Behaviour Analysis and Discrete Trial Training. More so, due to the unpredictable nature of children on the autism spectrum, it is proposed that future works could look at how ensemble machine learning models which utilize facial emotions, audio and body posture can be utilized to predict the emotional states of the children. There is the need for researchers to also devote effort towards collecting datasets of autistic children to be able develop relevant machine learning applications to suit the needs of the children and their caregivers.

The experiments presented in this thesis were only conducted in three schools in Greater Accra Region and Kumasi. This research needs to be extended to other parts of the country and more children can be added to the study in order to assess the effects of the robot intervention. Research can also focus on how factors such as age, IQ level and preexposure to technological devices can influence the categories of shapes, sizes and features of robots that would appeal most to children with autism.

REFERENCES

Adamson, L. (1995). *Communication development during infancy*. Madison, WI: Brown and Benchmark Publishers.

Agarwal, K. (2011). *Object Detection in Refrigerators using Tensor*.

Alnahdi, G. H. (2013). Single-subject designs in special education:advantages and limitations. *Journal of Research in Special Education Needs*.

Alonso-Martin, F., Malfaz, M., Sequeira, J., Gorostiza, J., & Salichs, M. (2013). A multimodal emotion detection system during human–robot interaction. *Sensors, 13(11)*.

Anantrasirichai, N., Achim, A., Morgan, J. E., & Erchova, I. (2013). SVM-based texture classification in optical coherence tomography. *In 2013 IEEE 10th International Symposium on Biomedical Imaging* (pp. 1332-1335). IEEE.

Anthony, J. (2011). Conceptualizing disability in Ghana: implications for EFA and inclusive education. *International Journal of Inclusive Education*.

Anzalone, S. M., Boucenna, S., I. S., & Chetouani, M. (2015). Evaluating the engagement with social robots. *International Journal of Social Robotics*.

APA. (2013). *American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders-5.*

Arlen, T. (2018). *Understanding the mAP Evaluation Metric for Object Detection.* Medium.

Aslam, S., Dertien, E. C., & van Dijk, J. (2019). Participatory Design of a Social Robot (So-bot) Toolkit for and with Adults with Autism. *4th RTD Conference: Design United* .

Ayodele, T. O. (2010). Types of machine learning algorithms. In *New Advances in machine learning.*

Bakare, M. O., & Munir, K. (2011). Autism spectrum disorders (ASD) in Africa: A perspective. *African Journal of Psychiatry* .

Bakare, M., & Munir, K. (2011). Autism spectrum disorders (ASD) in Africa: a perspective. *African Journal of Psychiatry*, 208-210.

Baker, E. W.-G., & Hubona, G. S. (2010). Cultural impacts on acceptance and adoption of information technology in a developing country. *Journal of Global Information Management (JGIM) 18.3* , 35-58.

Barron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review." Understanding other minds. *Perspectives from developmental cognitive neuroscience 2*.

Bi, Y., & Jeske, D. R. (2010). The efficiency of logistic regression compared to normal discriminant analysis under class-conditional classification noise. *Journal of Multivariate Analysis, 101(7)*, 1622-1637.

Bins, J., & Draper, B. (2001). Feature selection from huge feature sets. *Proceedings of the Eighth IEEE International Conference on Computer Vision, ICCV, 2*. IEEE.

Bishop, S. K., Moore, J. W., Dart, E. H., Radley, K., B. R., Barker, L. K., . . . Toche, C. (2019). Further investigation of increasing vocalizations of children with autism with a speech-generating device. *Journal of applied behavior analysis*.

Blacher, J., & Mink, I. T. (2004). *Interviewing family members and care providers: concepts, methodologies, and cultures*. Chichester: JohnWiley & Sons.

Boehm, M. (2016). *How is autism diagnosed?* New Mexico: Center for development and disability.

Brian, S., Admoni, H., & Matarić, M. (2012). Robots for use in Autism research. *Annual review of biomedical engineering 14*, 275-294.

Cabibihan, J., Hifza, J., Marcelo, A., & Aljunied, S. M. (2013). A survey on the roles and benefits of social robots for the therapy of children with autism. *5*(4 593618).

Cascio, M. A. (2015). Cross-cultural autism studies, neurodiversity, and conceptualizations of autism. *Cult. Med. Psychiatry 39*, 207–212 .

Castellano, G., Caridakis, G., Camurri, A., Karpouzis, K., Volpe, G., & Kollias, S. (2010). *Body gesture and facial expression analysis for automatic affect recognition." Blueprint for affective computing: A sourcebook.*

CDC. (2014). *Centers_for_Disease_Control_and_Prevention, Autism Spectrum disorder, data and statistics.*

CDC, C. f. (2016). *Summary of Autism Spectrum Disorder (ASD) Prevalence Studies.*

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Auriol, D., & Cox, A. (2003). Predicting language outcomes in infants with autism and pervasive developmental disorder. *International Journal of Languages and Communication disorders 38(3)*, 265-285.

Charmman, T. (2003). Why is joint attention a pivotal skill in autism? *Philos Trans R Soc Lond B Biol Sci. 358(1430)*, 315–324.

Chen, Z., Zhang, T., & Ouyang, C. (2018). End-to-End Airplane Detection Using Transfer Learning in Remote Sensing Images. *Remote Sens*.

Cheng, Y., Lu, F., & Zhang, X. (2018). Appearance-Based Gaze Estimation via Evaluation-Guided Asymmetric Regression. *Proceedings of the European Conference on Computer Vision (ECCV).*

Christensen, B. J., Van Naarden Braun, K., Bilder, D., J., C., Constantino, J., Daniels, J., . . . CDC. (2016). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years. *Errata, MMWR Surveill Summ. 65(SS-3)*, 1-23.

Cid, F., Prado, J. A., Bustos, P., & Nunez, P. (2013). A Real Time and Robust Facial Expression Recognition and Imitation approach for Affective Human-Robot Interaction Using Gabor filtering. *IEEE International Conference on Intelligent robots and Systems.* IEEE.

Coeckelbergh, M., Pop, C., Simut, R., Peca, A., Esteban, P., De Beir, A., . . . Vanderborght, B. (2015). Survey investigating ethical issues concerning Robot Enhanced Therapy for children with autism. *1st International Conference on Social Robots in Therapy and Education,*. Almere, The Netherlands.

Coeckelbergh, M., Pop, C., Simut, R., Peca, A., Pintea, S., David, D., & Vanderborght, B. (2016). A survey of expectations about the role of robots in robot-assisted therapy for children with ASD: ethical acceptability, trust, sociability, appearance, and attachment. *Science and engineering ethics*, *22*(1), 47-65.

Colton, M. B., Ricks, D. J., Goodrich, M. A., Dariush, B., Fujimura, K., & Fujiki, M. (2009). oward therapist-in-the-loop assistive robotics for children with autism and specific language impairment. *Autism 24, 25*.

Conti, D., Cattani, A., Di Nuovo, S., & Di Nuovo, A. (2015). A cross-cultural study of acceptance and use ofrobotics by future psychology practitioners. *24th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN).* IEEE.

Conti, D., Nuovo, S., Buono, S., & Nuovo, A. (2017). Robots in Education and Care of Children with Developmental Disabilities: A Study on Acceptance by Experienced and Future Professionals. *International Journal of Social Robotics*.

Crabbe, M., Standing, C., Standing, S., & Karjaluoto, H. (2009). An adoption model for mobile banking in Ghana. *International Journal of Mobile Communications*.

Crippa, A., Salvatore, C., Perego, P. F., & Nobile, M. (2015). Use of Machine Learning to Identify Children with Autismand Their Motor Abnormalities. *J Autism Dev Disord*.

Dai, J., Li, Y., He, K., & Sun, J. (2016). R-fcn: Object detection via region-based fully convolutional networks. *Advances in neural information processing systems,* (pp. 379-387).

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *In international Conference on computer vision & Pattern Recognition (CVPR'05)* . IEEE.

Dautenhahn, K., & Iain, W. (2004). Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition 12.1*, 1-35.

Dautenhan, K. (2003). Roles and functions of robots in human society: implications from research in autism therapy. *Robotica volume 21*, 443-452.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13(3)*, 319-340.

de Oliveira, F. A., Nobre, C. N., & Zarate, L. E. (2013). Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index–Case study of PETR4. *Expert Systems with Applications, 40(18)* , 7596-7606.

De Ruyter, B., Saini, P., Markopoulos, P., & . van Breemen, A. J. (2005). Assessing the Effects of Building Social Intelligence in a Robotic Interface for the Home. *Special Issue of IwC: social impact of emerging technologies 17(5)*, 522-541.

De Silva, P., Tadano, K., Saito, A., Lambacher, S., & Higashi, M. (2009). Therapeuticassisted robot for children with autism. *The IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Diehl, J. J., Schmitt, L. M., Villano, M., & Crowell, C. R. (2012). The clinical use of robots for individuals with autism spectrum disorders: A critical review. *Research in autism spectrum disorders, 6(1), 249-262*.

Dillon, A., & Morris, M. G. (1996). User acceptance of information technology: Theories and models. *Annual Review of Information Science and Technology 31*, 3-32.

Djam, X., Wajiga, G. M., Kimbi, Y. H., & N.V., B. (2011). A Fuzzy Expert System for the Management of Malaria. *International Journal of Pure and Applied Sciences and Technology 5(2).*

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and autonomous systems 42.3*, 177-190.

Duquette, A., Michaud, F., & Mercier, H. (2008). Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism. *Autonomous robots 24(2)*, 147–157.

Dyches, T., Wilder, L., Sudweeks, R., Obiakor, F., & Algozzine, B. (2004). Multicultural issues in autism. *Journal of Autism and Developmental Disorders, 34*, 211-222.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the. *Journal of personality and social psychology*.

Elfenbein, A. H., & Ambady, N. (2002). On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis. *Psychological Bulletin 128(2).*

Ennis-Cole, D., Durodoye, B. A., & Harris, H. L. (2013). The Impact of Culture on Autism Diagnosis and Treatment:Considerations for Counselors and other Professionals. *The Family Journal: Counseling and Therapy for Couples and Family 00(0)*, 1-9.

Fang, Z., & Zheng, J. (2019). Object Grasping of Humanoid Robot Based on YOLO. *Advances in Computer Graphics: 36th Computer Graphics International Conference, CGI.* Calgary.

Feil-seifer, D., & Viterbi, U. (2009). Development of socially assistive robots for children with autism spectrum disorders.

Feng, Y., Jia, Q., Chu, M., & Wei, W. (2017). Engagement Evaluation for Autism Intervention by Robots Based on Dynamic Bayesian Network and Expert Elicitation. *IEEE Access, 5*.

Fernandez Robles, L. (2016). *Object recognition techniques in real applications.* University of Groningen.

Flandorfer, P. (2012). Population Ageing and Socially Assistive Robots for Elderly Persons: The Importance of Sociodemographic Factors for User Acceptance. *International Journal of Population Research*.

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: an update. *J Autism Dev, 33*, 365 – 382.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems 42(3-4)*, 143-166.

Forlizzi, J., & Bartneck, C. (2004). A Design-Centred Framework for Social HumanRobot Interaction. *Proceedings of the Ro-Man2004*, (pp. pp. 591-594). Kurashiki.

Fuentes, J., Muideen, B., Kerim, M., Patricia, A., Naoufel, G., Özgür, Ö., & Marcos, M. (2012). *Autism Spectrum Disorders.* Geneva: International Association for Child and Adolescent Psychiatry and Allied Professions.

Gernsbacher, M. A., Stevenson, J. L., Khandakar, S., & Goldsmith, H. H. (2008). Why Does Joint Attention Look Atypical in Autism? *Child Development Perspectives 2(1)*, 38–45.

Ghahramani, Z. (2004). Unsupervised Learning. *Advanced lectures on machine learning*.

Giorgan, G., & Ploeger, P. G. (2011). Facial Expression Recognition for Domestic Service Robots. *Robot Soccer World Cup*.

Girshick, R. (2015). Fast r-cnn. *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448). IEEE.

Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., & Agrawal, A. (2017). Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials, 157*, 322-330.

Gottesman, O., Johansson, F., Meier, J., Dent, J., Lee, D., Srinivasan, S., & ... & Yao, J. (2018). *arXiv preprint arXiv:1805.12298*.

Gu, J., Wang, Z., Kuen, J., M. L., Shahroudy, A., Shuai, B., & ... & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition 77*, 354377.

Heerink, M., Kröse, B. ,., & Wielinga, B. (2006). The Influence of a Robot's Social Abilities on Acceptance by Elderly Users. *Robot and Human Interactive Communication.* IEEE ROMAN.

Henderson, P., & Ferrari, V. (2016). End-to-end training of object class detectors for mean average precision. *Asian Conference on Computer Vision.* Cham: Springer.

Huijnen, C. A., Lexis, M. A., & Witte, L. P. (2017). Robots as New Tools in Therapy and Education for Children with Autism. *International Journal of Neurorehabilitation*.

Hutson, S., Lim, S., Bently, P., Bianchi-Berthouze, N., & Bowling, A. (2011). Investigating the Suitability of Social Robots for the Wellbeing of the Elderly. *Affective Computing and Intelligent Interaction. ACII 2011*. Springer, Berlin.

Ingersoll, B. (2007). Teaching Imitation to Children with Autism: A focus on social reciprocity. *SLP- ABA 2(3)*.

Ingersoll, B. (2008). The Social Role of Imitation in Autism:Impliations for the treatment of imitation deficits. *21(2),* 107–119.

Jain, P., & Rekha. (2017). Impact of Demographic Factors: Technology Adoption in Agriculure. *SCMS Journal of Indian Management*.

Jiang, Y., Chen, D., & Lai, F. (2010). Technological-Personal-Environmental (TPE) Framework: A Conceptual Model for Technology Acceptance at the Individual Level. *International Information Management Association* .

Jones, E. A., & Carr, E. G. (2004). Joint Attention in Children With Autism. *Focus on Autism and other Developmental Disabilities 19(1)*.

Kahou, S., Michalski, V., & Konda, K. (2015). Recurrent neural networks for emotion recognition in video. *Proceedings of the ACM on International Conference on Multimodal Interaction.* Seattle: ACM.

Kim, E., Berkovits, L., Bernier, E., Leyzberg, D., Shic, F., Paul, R., & Scassellati. B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. . *Journal of Autism and Developmental Disorders*, 1038-1049.

Kim, J. C., Azzi, P., Jeon, M., Howard, A. M., & Park, C. H. (2017). Audio-based emotion estimation for interactive robotic therapy for children with autism spectrum disorder. *14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI).* IEEE.

Kloo, D., Perner, J., & Gritzer, T. (2010). *Object-based set-shifting in preschoolers: Relations to Theory of Mind.* New York: In B. W. Sokol, U. Müller, J. I. M. Carpendale, A. R. Young, & G. Iarocci (Eds.), Oxford University Press.

Kozima, H., & Nakagawa, C. (2006). Interactive robots as facilitators of childrens social development. *Mobile robots: Towards new applications. Intech*.

Kozima, H., Cocoro, N., & Yuriko, Y. (2007). Children–robot interaction: a pilot study in autism therapy. . *Progress in brain research (164)*, 385-400.

KPMG. (2016). *2016's new breed of social robots is ready to enter your world.*

Krupa, N., Anantharam, K., Sanker, M. D., & Sagar, J. V. (2016). Recognition of emotions in autistic children using physiological signals. *Health and Technology 6(2)*, 137–147.

Learned-Miller, E. G. (2014). *Introduction to supervised learning.* Department of Computer Science, University of Massachusetts.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature 521*.

Lee, J., Hiroki, T., Chikara, N., Goro, O., & Dimitar, S. (2012). Which Robot Features Can Stimulate Better Responses from Children with Autism in Robot-Assisted Therapy? . *International Journal of Advanced Robotic Systems, 9(3).*

Leo, M., Coco, M. D., Carcagni, P., & Distante, C. (2015). Automatic Emotion Recognition in Robot-Children Interaction for ASD treatment. *Proceedings of the IEEE International Conference on Computer Vision Workshops.* IEEE.

Levi, G., & Hassner, T. (2015). Emotion Recognition in the Wild via Convolutional Neural Networks and Mapped Binary Patterns. *ACM International Conference on Multimodal Interaction (ICMI).* Seattle.

Li, J., Chen, X., Hovy, E., & Jurafsky, D. (2015). Visualizing and understanding neural models in nlp. *arXiv preprint arXiv:1506.01066.*

Li, S., & Deng, W. (2018). Deep Facial Expression Recognition: A Survey. *arXiv preprint arXiv:1804.08348* .

Liang, X., Wang, T., Yang, L., & Xing, E. (2018). Cirl: Controllable imitative reinforcement learning for vision-based self-driving. *Proceedings of the European Conference on Computer Vision (ECCV)* , (pp. 584-599).

Libin, A., & Libin, E. (2004). Person-robot interactions from the robopsychologists' point of view: the robotic psychology and robotherapy approach. *Proc. IEEE 92*, 1789–1803.

Lisa, S., Mundy, P., & Mastergeorge, A. M. (2015). Joint Attention in Preschool Children: Is it a Meaningful Measure? *International Journal of School and congnitive psychology 2:120*.

Littlewort, G., Bartlett, M., Fasel, I., Chenu, J., & Kanda, T. (2004). Towards social robots: Automatic evaluation of human-robot interaction by face detection and expression classification. *Advances in Neural Information Processing Systems*.

Liu, C., Conn, K., Sarkar, N., & Stone, W. (2008). Online Affect Detection and Robot Behavior Adaptation for Intervention of Children With Autism. *IEEETransactions on Robotics, 24(4)*.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. *European conference on computer vision* . Charm: Springer.

Liu, Z., Wu, M., Cao, W., & Chen, L. (2017). A Facial Expression Emotion Recognition Based Human-robot Interaction System. *IEEE/CAA Journal of Automatica Sinica*.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. *ICCV.*

Lowry, L. (2016). *Imitation with Children on the Autism Spectrum: More Than Just a Game of Copycat.* The Hanen Center.

MacDorman, K. F., Minato, T., Shimada, M., Itakura, S., Cowley, S. J., & Ishiguro, H. (2005). Assessing human likeness by eye contact in an android testbed. *Proceedings of the XXVII Annual Meeting of the Cognitive Science Society.* Italy.

Mandell, D. S., & Novak, M. (2005). The role of culture in families' treatment decisions for children with autism spectrum disorder. *Mental retardation and developmental disabilities research reviews*.

Mandell, D., Novak, M., & Zubritsky, C. (2007). Factors associated with age of diagnosis among children with autism spectrum disorders. *Journal of Autism Development Disorder 37(10)*, 1902-1912.

Mayya, V., Pai, R., & Pai, M. (2016). Automatic facial expression recognition using DCNN. *Procedia Computer Science, 93*, 453-461.

MCA, M. (2012). *Addressing the needs of individuals with autism spectrum disorder in Maryland.* Maryland.

McColl, D., Hong, A., Hatakeyama, N., Nejat, G., & Benhabib, B. (2016). A Survey of Autonomous Human Affect Detection Methods. *Intelligent Robot Systems*.

Meyes, R., Tercan, H., Roggendorf, S., Thiele, T., Büscher, C., Obdenbusch, M., & ... & Meisen, T. (2017). Motion planning for industrial robots using reinforcement learning. . *Procedia CIRP(63)*, 107-112.

Michaud, F., & Caron, S. (2002). Roball, the rolling robot. *Autonomous Robots 12*, 211222.

Michaud, F., Tamier, S., Duquette, A., Mercier, H., Michel, L., Larouche, H., & Larose, F. (2007). Assistive Technologies and Child-Robot Interaction. *AAAI spring symposium on multidisciplinary collaboration for socially assistive robotics*.

Miikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., . . . Hodjat, B. (2017). Evolving deep neural networks. *arXiv:1703.00548 [cs.NE]*.

Mitzner, T., Boron, J., Fausset, C., Adams, A., Charness, N., Czaja, S., . . . Sharit, J. (2010). Older Adults Talk Technology: Technology Usage and Attitudes. *Comput Human Behaviour*, *26(6)*, 1710-1721.

Mollahosseini, A., Chan, D., & Mahoor, M. H. (2015). Going Deeper in Facial Expression Recognition using Deep Neural Networks. *IEEE winter conference on applications of computer vision (WACV).* IEEE.

Mori, M. (1970). The uncanny valley". *Energy 7(4)*, 33-35.

Moslehi, H. R. (2011). *Design and development of a fuzzy logic operated mocrocontroller based smart motorized wheelchair*.

Mubin, O., Stevens, C. J., & Dong, J.-J. (2013). A review of the applicability of robots in education. *Technology for Education and Learning*.

Muda, H. M., Saad, P., & Othman, R. M. (2011). Remote protein homology detection and fold recognition using two-layer support vector machine classifiers. *Computers in biology and medicine 41(8)*, 687-699.

Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*.

Nair, I., & Das, V. M. (2012). Using Technology Acceptance Model to assess teachers' attitude towards use of technology as teaching tool: A SEM Approach. *International Journal of Computer Applications, 42(2)*.

NaoRobotics. (2012). *Nao Robotics.* Retrieved February 2018, 26, from Twitter: https://twitter.com/naorobots

Naqa, I. E., & Murphy, M. J. (2015). What Is Machine Learning? In *Machine Learning in Radiation Oncology*. Cham: Springer.

Nazemizadeh, M., Taheri, M., & Nazeri, S. (2014). he application of fuzzy-logic method to control of robots: a review study. *International Journal of Mechanical Engineering and Robotics* .

NIH. (2015). *National Institute of Mental Health, Autism Spectrum disorder.* National Institute of Mental Health.

Okeke, I. (2016). The Right to Free Education of Autistic Nigerian Children.

Owoseni, A. T., & Ogundahunsi, I. O. (2016). Mobile-Based Fuzzy Expert System for Diagnosing Malaria (MFES). *I.J. Information Engineering and Electronic Business*.

Padilla, R., Netto, S. L., & da Silva, E. A. (2020). A Survey on Performance Metrics for Object-Detection Algorithms. *International Conference on Systems, Signals and Image Processing (IWSSIP)* , (pp. 237-242).

Pedersen, T. (2018). *Theory of Mind.* Retrieved September 27, 2018, from

PschCentral.com: https://psychcentral.com/encyclopedia/theory-of-mind/ Piaget, J.

(1974). *To Understand is to Invent.* . NY:: Basic Books.

Pitten, K. (2008). How cultural values influence diagnosis,treatment and the welfare of families with an autistic child. *Insight: rivier academic journal, volume 4, number 1*.

Quiros-Ramirez, M. A., & Onisawa, T. (2015). Considering cross-cultural context in the automatic recognition of emotions. *nternational Journal of Machine Learning and Cybernetics*.

Radovan, M., & Kristl, N. (2017). Acceptance of Technology and its Impact on Teacher's Activities in Virtual Classroom. *TOJET: The Turkish Online Journal of Educational Technology*.

Rastall, E. (2016). *Understanding Autism and theory of the mind.* Seattle Children's hospital.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition.*

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 91-99.

RIRSGH. (2010). *Rural Integrated Relief Service-Ghana, The Autism Project for Ghana.*

Ritschel, H. (2018). *Socially-Aware Reinforcement Learning for Personalized HumanRobot Interactions.*

Robins, B. (2005). A Humanoid Robot as Assistive Technology for Encouraging Social Interaction Skills in Children with Autism.

Robins, B., Dautenhahn, K., & Dickerson, P. (2009). From Isolation to Communication: A Case Study Evaluation of Robot Assisted Play for Children with Autism with a Minimally Expressive Humanoid Robot. *Second International Conferences on Advances in Computer-Human Interactions (pp. 205-211).* Cancun: IEEE.

Robins, B., Dautenhahn, K., & Dubowski, J. (2006). “Does appearance matter in the interaction of children with autism with a humanoid robot? *Interaction studies 7(3)*, 479-512.

Robins, B., Dautenhahn, K., te Boekhorst, R., & Billard, A. (2004). Effects of Repeated Exposure to a Humanoid Robot on Children with Autism. In *Designing a More Inclusive World. Springer, London* (pp. 225-236). London: Springer.

Robins, B., Dautenhahn, K., Te Boekhorst, R., & Billard, A. (2005). Robotic assistants in therapy and education of children with autism:can a small humanoid robot help encourage social interaction skills? *Univ Access Inf Soc 4*, 105–120.

Robins, B., Dickerson, P., Stribling, P., & Dautenhahn, K. (2004). Robot-mediated joint attention in children with autism: A case study in robot-human interaction. *Interaction studies 5(2)*, 161-198.

Roth, A. V. (2007). Applications of emperical science in manufacturing and service operations. *Manufacturing & Service Operations Management*.

Rouast, P. V., Adam, M. T., & Chiong, R. (2019). Deep Learning for Human Affect Recognition:Insights and New Developments. *IEEE Transactions on Affective Computing*.

Rudovic, O., Lee, J., Mascarell-Maricic, L., Schuller, B. W., & Picard, R. W. (2017). Measuring Engagement in Robot-Assisted Autism Therapy: A Cross-Cultural Study. *Frontiers in Robots and AI*.

Rudovic, O., Utsumi, Y., Lee, J., Hernandez, J., Ferrer, E. C., Schuller, B., & Picard, R. W. (2018). CultureNet: A Deep Learning Approach for Engagement Intensity Estimation from Face Images of Children with Autism. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* . IEEE.

Ruparelia, K. A. (2016). Autism SpectrumDisorders in Africa: Current challenges in identification, Assessment and treatment: A report on the International Child Neurology Association Meeting on ASD in Africa. *Journal of Child Neurology*.

Sam, G., & Naglieri, J. A. (2010). *Autism Spectrum Rating Scales: ASRS.* North Tonawanda, NY: Multi-Health Systems.

Samadi, S. A., & McConkey, R. (2011). Autism in Developing Countries: Lessons from Iran. *Autism research and treatment*.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE.

Sanua, V. (1984). Is infantile autism a universal phenomenon? An open question. *Int J Soc Psychiatryn 30(3)*, 163-177.

Scassellati, B. (2005). Quantitative metrics of social response for autism diagnosis. *In IEEE International Workshop on Robots and Human Interactive Communication (ROMAN)*, 585–590.

Scassellati, B., Admoni, H., & Mataric, M. (2012). Robots for use in autism research. *Scassellati, Brian, Henny Admoni, and Maja Matarić. "Robots for use in autism research." Annual review of biomedical engineering 14*, 275-294.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of autism and developmental disorders, 10(1)*, 91-103.

Shamsuddin, S., Yussof, H., Ismail, L. I., Mohamed, S., Hanapiah, F. A., & Zahari, N. I. (2012). Humanoid robot NAO interacting with autistic children of moderately impaired intelligence to augment communication skills. *Procedia Engineering, 41*, 1533-1538.

Shan, C., Gong, S., & McOwan, P. W. (2009). Facial expression recognition based on local binary patterns: A comprehensive study. *Image and Vision Computing*.

Soviany, P., & Ionescu, R. T. (2018, September). Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction. In *2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)* (pp. 209-214).

Smitha, K. G., & Vinod, A. P. (2015). Facial emotion recognition system for autistic children: a feasible study based on FPGA implementation. *Medical Biological Engineering Computing*.

So, W. C., Wong, M. K., & Lam, K. Y. (2016). Social and communication skills predict imitation abilities in children with autism. *Frontiers in Education 1(3)*.

Srite, M., & Karahanna, E. (2006). The role of espoused national cultural values in technology acceptance. *MIS quarterly* , 679-704.

Stanton, C., Kahn, P. J., Severson, R., Ruckert, J., & Gill, B. (2008). Robotic animals might aid in the social development of children with autism. *Proc. 3rd ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI '08).* New York.

Stein, G., Chen, B., Wu, A. S., & Hua, K. A. (2005). Stein, G., Chen, B., Wu, A. S.,Decision tree classifier for network intrusion detection with GA-based feature selection. *Proceedings of the 43rd annual Southeast regional conference-Volume 2 (*, (pp. 136-141).

Stone, W. L., Ousley, O. Y., & Littleford, C. D. (1997). Motor imitation in young children with autism: What's the object? *Journal of Abnormal Child Psychology, 25*, 475–485.

Sun, H., & Zhang, P. (2006). The role of moderating factors in user technology acceptance. *International Journal of Human Computer Studies 64(2)*.

Sutton, R. S., & Barto, A. G. (2015). *Reinforcement Learning: An Introduction (In Progress).* MIT Press.

Syamimi, S., Hanafiah, Y., Salina, M., & Fazah, A. H. (2014). Design and Ethical Concerns in Robotic Adjunct Therapy Protocols for Children with Autism. *Procedia Computer Science 42* , 9-16.

Thabtah, F. (2017). Machine learning in autistic spectrum disorder behavioralresearch: A review and ways forward. *Informatics for Health and Social Care*.

Tincani, M., Travers, J., & Boutot, A. (2009). Race, culture, and autism spectrum disorder: understanding the role of diversity in successful educational interventions. *Res. Pract. Pers. Sev. Disabil. 34*, 81–90.

Torrey, L., & Shavlik, J. (2010). Transfer learning. In Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. *IGI Global*.

.

Ueyama, Y. (2015). Ueyama, Yuki. "A bayesian model of the uncanny valley effect for explaining the effects of therapeutic robots in autism spectrum disorder. *PloS one 10(9)*.

United_Nations, U. N. (2019). *World Autism Awareness Day 2nd April Report.*

Valadão, C. T., Goulart, C., Rivera, H., Caldeira, E., Filho, T. F., Frizera-Neto, A., & Carelli, R. (2016). Analysis of the use of a robot to improve social skills in children with autism spectrum disorder. *Research on Biomedical Engineering 32(2)*.

Valentina, P. (2017). Social robotics to help children with autism in their interactions through imitation. *REM - Research on Education and Media, Volume 9*.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46 (2),* 186–204.

Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly 27(3)*, 425-478.

Wainer, J., Kerstin, D., Ben, R., & Farshid, A. (2010). Collaborating with Kaspar: Using an autonomous humanoid robot to foster cooperative dyadic play among children with autism. *10th IEEE-RAS International Conference on Humanoid Robots.* IEEE.

Wainer, J., Robins, B., Amirabdollahian, F., & Dautenhahn, K. (2014). Using the humanoid robot KASPAR to autonomously play triadic games and facilitate collaborative play among children with autism. *IEEE Transactions on Autonomous Mental Development, 6(3),* 183-199.

Wang, L., Li, R., Shi, H., S. J., Zhao, L., S. H., . . . Tandianus, B. (2019). Multi-Channel Convolutional Neural Network Based 3D Object Detection for Indoor Robot Environmental Perception. *Sensors, 19(4)*.

Wang, Y., Wang, C., & Zhang, H. (2018). Combining a single shot multibox detector with transfer learning for ship detection using sentinel-1 SAR images. *Remote sensing letters, 9(8),* 780-788.

Wany_Robotics. (2008). *Pekee Mobile Robot Platform.* Retrieved Februaary 26, 2018, from Wany Robotics: http://www.wany.fr/

Weiss, A. (2016). Creating Service Robots for and with People: A User-centered Reflection on the Interdisciplinary Research Field of Human-Robot Interaction. *15th Annual STS Conference Graz, Critical Issues in Science, Technology, and Society Studies.*

Weiss, A., Wurhofer, D., Bernhaupt, R., Altmaninger, M., & Tscheligi, M. (2010). A Methodological Adaptation for Heuristic Evaluation of HRI. *IEEE RO-MAN*.

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big data, 3(1), 9*.

Wenbo, L., Ming, L., & Li, Y. (2016). Identifying Children with Autism Spectrum Disorder Based on Their Face Processing Abnormality: A Machine Learning Framework. *Autism Research 9.8 (2016)*, 888-898.

Werry, I., Dautenhahn, K., & Harwin, W. (2001). Investigating a robot as a therapy partner for children. *In Proceedings of the European conference for the advancement of Assistive Technology*, 3-6.

Wimmer, M., MacDonald, B., Jayamuni, D., & Yadav, A. (2008). Facial Expression Recognition for Human-robot Interaction–A Prototype. *Robot Vision 4931*, 139–152.

Wong, G. K. (2015). Wong, Gary KW. "Understanding technology acceptance in preservice teachers of primary mathematics in Hong Kong. *Australasian Journal of Educational Technology 31(6)*.

Wootaek, L., Daeyoung, J., & Taejin, L. (2016). Speech Emotion Recognition using Convolutional and Recurrent Neural Networks. *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA.* IEEE.

Wu, P. F. (2009). User Acceptance of Emergency Alert Technology: A Case Study . *Proceedings of the 6th International ISCRAM Conference* . Gothenburg, Sweden.

Xu, Q., Jamie, N., Tan, O., Huang, Z., Tay, B., & Park, T. (2015). Methodological Issues in Scenario-Based Evaluation of Human–Robot Interaction. *International Journal of Social Robotics*.

YamashitaRikiya, Nishio, M., Do, R. K., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging 9*, 611–629.

Yang, B. S., Oh, M. S., & Tan, A. C. (2009). Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference. *Expert Systems with Applications, 36(2)*, 1840-1849.

Yogesh, K. D., Nripendra, P. R., Hsin, C., & Williams, M. D. (2011). A Meta-analysis of the Unified Theory of Acceptance and and Use of Technology (UTAUT). In *IFIP Advances in Information and Communication Technology 366.* Berlin, Heidelberg : Springer.

Young, N. D., & Jean, E. (2019). *Autism Spectrum Disorder. Creating Compassionate Classrooms: Understanding the Continuum of Disabilities and Effective Educational Interventions.*

Yu, Z., & Zhang, C. (2015). Image based static facial expression recognition with multiple deep network learning. *International Conference on Multimodal Interaction.* ACM.

Yun, J., Rong, C., Xiaoyan, K., Kangkang, C., Zuhong, L., & Herskovits, E. H. (2010). Predictive models of autism spectrum disorder based on brain regional cortical thickness. *Neuroimage*.

Zalah, I. (2018). *Factors That Influence Saudi Secondary Teachers' Acceptance and Use of E-Learning Technologies.*

Zhang, S. X., & Gales, M. J. (2012). Structured SVMs for automatic speech recognition. *IEEE Transactions on Audio, Speech, and Language Processing, 21(3)*, 544-555.

Zhang, X., Sugan, Y., Fritz, M., & Bulling, A. (2015). Appearance-Based Gaze Estimation in the Wild. *IEEE Conference on Computer Vision and Pattern Recognition.* IEEE.

Zhi, R., Flierl, M., Ruan, Q., & Kleijn, W. B. (2011). Graph-preserving sparse nonnegative matrix factorization with application to facial expression recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B.*

Zhou, Y., Yu, F., & Duong, T. (2014). Multiparametric MRI Characterization and Prediction inAutism Spectrum Disorder Using Graph Theory andMachine Learning. *PLoS ONE 9(6).*

Zimmerman, A. (2008). *Autism: Current theories and evidence.* Springer Science & Business Media.

PUBLICATIONS

This thesis has produced the following publications:

1. Mensah, R. O., Hayfron-Acquah, J. B., & Asante, M (2019). "RoCA: A humanoid robotic assistive technology for children with autism. Accepted for publication in *International Journal of Computer Engineering and Information Technology (IJCEIT) 11(8).*

2. Mensah, R. O., Hayfron -Acquah, J. B., & Asante, M (2019). Longitudinal study of interactions among Ghanaian autistic children and Rosye, a humanoid robot. *International Journal of Computer Techniques (IJCT)* 6(3)

3. Mensah, R.O., & Hayfron-Acquah, J. B. (2018).Social Robotic Systems in Autism Therapy: Survey and Design Considerations. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)* 6(5).

4. Mensah, R.O., & Hayfron -Acquah, J. B. (2018). Preliminary Ob servations from Interactions among Ghanaian Autistic Children and Rosye, a Humanoid Robotic Assistive Technology. *International Journal of Advances in Scientific Research and Engineering*, 4.

APPENDIX

A. EXPERIMENTAL DATASHEETS

Abbreviations

LHU – Left Hand Up
LHD – Left Hand Down
RHU – Right Hand Up
RHD – Right Hand Down
BHU – Both Hands Up
BHD – Both Hand Down

1 i. Data sheet for the scores of James in the imitation (Im) tasks

Session	LHU	LHD	RHU	RHD	BHU	BHD	Total Im. Score per session
S1	1	1	1	1	1	1	6
S2	1	1	1	1	1	1	6
S3	1	1	1	1	1	1	6
S4	1	1	1	1	1	1	6
S5	1	1	1	1	1	1	6
S6	1	1	1	1	1	1	6
S7	1	1	0	0	0	0	2
S8	1	1	1	1	1	1	6
Overall imitation score for the 8 sessions							**44**

ii. Data sheet for the scores of

James in the General Activity (GA) tasks

Session	Respond to name	What is your name?	Take the ball	Give me the ball	How are you?	Total GA. Score per session
S1	1	1	1	1	1	5
S2	1	1	1	1	0	4
S3	1	0	1	1	1	4
S4	1	1	1	1	1	5
S5	1	0	1	0	1	3
S6	1	0	1	1	1	4
S7	1	0	1	1	1	4
S8	0	0	1	1	1	3
Overall Activity score for 8 sessions						**32**

John

2 i. Data sheet for the scores of John in the imitation (Im) tasks

Session	LHU	LHD	RHU	RHD	BHU	BHD	Total Im. Score per session
S1	1	1	1	1	1	1	6
S2	1	1	1	1	1	1	6
S3	1	1	1	1	1	1	6
S4	1	1	1	1	1	1	6
S5	1	1	1	1	1	1	6
S6	1	1	1	1	1	1	6
S7	1	1	1	1	1	1	6
Overall imitation score for 7 sessions							**42**

ii. Data sheet for the scores of in the General Activity (GA) tasks

John

Session	Respond to name	What is your name?	Take the ball	Give me the ball	How are you?	Total GA. Score per session
S1	1	1	1	1	1	5
S2	1	1	1	0	1	4
S3	1	1	1	1	1	5
S4	1	1	1	1	1	5
S5	1	1	1	1	1	5
S6	1	0	1	1	1	4
S7	1	1	1	1	1	5
Overall Activity score for the 7 sessions						**33**

Ama

3 i. Data sheet for the scores of Ama in the imitation (Im) tasks

Session	LHU	LHD	RHU	RHD	BHU	BHD	Total Im. Score per session
S1	0	0	0	0	0	0	0
S2	0	0	0	0	0	0	0
S3	0	0	0	0	0	0	0
S4	0	0	0	0	1	1	2
S5	0	0	0	0	1	1	2
S6	1	1	0	0	1	1	4

ii. Data sheet for the scores of in the General Activity (GA) tasks

S7	0	0	0	0	1	1	2
Overall imitation score for 7 sessions							**10**

Ama

Session	Respond to name	What is your name?	Take the ball	Give me the ball	How are you?	Total GA Score per session
S1	0	0	1	1	1	3
S2	1	0	1	0	1	3
S3	1	0	1	1	1	4
S4	1	0	1	1	0	3
S5	1	0	1	1	1	4
S6	1	0	1	1	1	4
S7	1	1	1	1	1	5
Overall Activity score for the 7 sessions						**26**

Kofi

4 i. Data sheet for the scores of Kofi in the imitation (Im) tasks

ii. Data sheet for the scores of   in the General Activity (GA) tasks

Session	LHU	LHD	RHU	RHD	BHU	BHD	Total Im. Score per session
S1	1	1	1	1	1	1	6
S2	0	0	1	1	1	1	4
S3	1	1	0	0	0	0	2
S4	1	1	1	1	1	1	6
S5	0	0	1	1	1	1	4
S6	0	0	1	1	1	1	4
S7	0	0	1	1	0	0	2
Overall imitation score for 7 sessions							**28**

ii. Data sheet for the scores of in the General Activity (GA) tasks Kofi

Session	Respond to name	What is your name?	Take the ball	Give me the ball	How are you?	Total GA Score per session
S1	1	1	1	1	1	5
S2	1	0	1	1	0	3
S3	1	0	1	0	1	3
S4	1	0	1	1	1	4
S5	1	0	1	0	1	3
S6	0	0	1	1	1	3
S7	0	1	1	0	0	2
Overall Activity score for the 7 sessions						**23**

Akwasi

5 i. Data sheet for the scores of Akwasi in the imitation (Im) tasks

Session	LHU	LHD	RHU	RHD	BHU	BHD	Total Im. Score per session
S1	0	0	0	0	0	0	0
S2	0	0	0	0	0	0	0
S3	0	0	0	0	0	0	0
S4	0	0	0	0	0	0	0
S5	0	0	0	0	0	0	0
Overall imitation score for 5 sessions							**0**

ii. Data sheet for the scores of

Akwasi in the General Activity (GA) tasks

Session	Respond to name	What is your name?	Take the ball	Give me the ball	How are you?	Total GA Score per session
S1	1	1	1	1	1	5
S2	1	0	1	0	1	3
S3	0	0	0	0	0	0
S4	0	0	0	0	0	0
S5	0	0	0	0	0	0
Overall Joint attention score for the 5 sessions						**8**

Yaw

6 i. Data sheet for the scores of Yaw in the imitation (Im) tasks

Session	LHU	LHD	RHU	RHD	BHU	BHD	Total Im. Score per session
S1	1	0	0	0	0	0	1
S2	1	1	0	0	0	0	2
S3	0	0	1	1	0	0	2
S4	0	0	0	0	0	0	0
S5	0	0	0	0	0	0	0
S6	0	0	0	0	0	0	0

ii. Data sheet for the scores of

S7	0	0	0	0	1	1	2
S8	0	0	0	0	0	0	0
Overall imitation score for 8 sessions							**7**

Yaw in the General Activity (GA) tasks

Session	Respond to name	What is your name?	Take the ball	Give me the ball	How are you?	Total GA Score per session
S1	1	0	1	0	1	3
S2	1	0	1	1	0	3
S3	1	0	0	0	0	1
S4	1	0	0	0	0	1
S5	0	0	0	0	0	0
S6	1	0	1	0	0	2
S7	1	0	0	0	0	1
S8	0	0	0	0	0	0
Overall Activity score for the 8 sessions						**11**

7. Scores of the number of times each child touched the robot over a maximum period of eight sessions

Session	James	John	Ama	Kofi	Akwasi	Yaw	Afia
S1	0	23	18	6	32	29	0
S2	0	12	4	19	0	18	0

ii. Data sheet for the scores of

S3	4	18	9	6	1	10	0
S4	4	13	1	14	2	3	0
S5	24	9	9	5	2	5	0
S6	63	49	3	35	x	1	0
S7	30	1	1	8	x	2	0
S8	17	X	x	x	x	0	0
Total number of times child touched the robot	**142**	**125**	**45**	**93**	**37**	**68**	**0**

x in the table means child was absent for that session

8. Comparison of the children’s responses to both tasks (imitation and joint

Child	total score for imitation tasks	total score for joint attention tasks	total number of times child touched the robot
James	44	32	142
John	42	33	125
Ama	10	26	45
Kofi	28	23	93
Akwasi	0	8	37
Yaw	7	11	68
Afia	0	0	0
Total	**131**	**133**	**510**

9. Total scores for each imitation task per session

Session	LHU	LHD	RHU	RHD	BHU	BHD	**total imitation score attained by all the children per session**
S1	4	3	3	3	3	3	19
S2	3	3	3	3	3	3	18
S3	3	3	3	3	2	2	16
S4	3	3	3	3	4	4	20
S5	2	2	3	3	4	4	18
S6	3	3	3	3	4	4	20
S7	2	2	2	2	3	3	14
S8	1	1	1	1	1	1	6
Total imitation score per task for all eight sessions	21	20	21	21	24	24	**131**

attention) and the number of times each child touched the robot

10. Total scores for each joint attention task per session

Session	Respond to name	What is your name?	Take the ball	Give me the ball	How are you?	**total joint attention score attained by all the children per session**
S1	5	4	6	5	6	26
S2	6	2	6	3	3	20
S3	5	1	4	3	4	17
S4	5	2	4	4	3	18
S5	4	1	4	2	4	15
S6	4	0	5	4	4	17
S7	4	3	4	3	3	17
S8	0	0	1	1	1	3
Total imitation score per task for all eight sessions	33	13	34	25	28	**133**

Overview of the parts of the robot touched by the children

	Parts of the robot							
Session(S)	Hand	Midsection	Head	Mouth	Neck	Shoulder	Eye	**Total**
S1	44	10	16	10	5	18	5	108
S2	21	6	11	2	1	7	5	53
S3	25	8	11	1	2	0	1	48
S4	20	12	5	0	0	0	0	37
S5	35	15	3	0	1	0	0	54
S6	75	33	16	11	16	0	0	151
S7	12	15	8	5	2	0	0	42

S8	8	5	2	0	2	0	0	17
Total	240	104	72	29	29	25	11	510

B. Labelmap.pbtxt details

```
item {     id: 1
name: 'Neutral'
}

item      {
id: 2
   name:
'Happy' } item {
id: 3     name:
'Sad'
}

item {
   id: 4
   name: 'Angry'
}

item {     id: 5
name: 'Fear'
}

item {     id: 6
name: 'Surprised'
}
```

C. Pipeline.config file details

```
model {   ssd {
num_classes: 6
image_resizer {
fixed_shape_resizer {
height: 300         width:
300
      }
    }
    feature_extractor {
type: "ssd_mobilenet_v2"
depth_multiplier: 1.0
min_depth: 16
     conv_hyperparams              {
regularizer {
        l2_regularizer {           weight:
3.9999998989515007e-05
        }
}
       initializer {
truncated_normal_initializer {
mean: 0.0            stddev:
0.029999999329447746
        }
       }
       activation: RELU_6
batch_norm {          decay:
0.9997000098228455
center: true
        scale: true
```

```
        epsilon: 0.0010000000474974513
        train: true
      }
    }
    use_depthwise: true
  }
  box_coder {
faster_rcnn_box_coder {
y_scale: 10.0        x_scale:
10.0        height_scale: 5.0
      width_scale: 5.0
    }
  }
  matcher {      argmax_matcher {
matched_threshold: 0.5
unmatched_threshold: 0.5
ignore_thresholds: false
negatives_lower_than_unmatched: true
      force_match_for_each_row: true
    }
  }
  similarity_calculator {
    iou_similarity {
    }
  }
  box_predictor                              {
convolutional_box_predictor {
      conv_hyperparams                {
regularizer {
        l2_regularizer {             weight:
3.9999998989515007e-05
        }
}
        initializer {
truncated_normal_initializer {
mean: 0.0
          stddev: 0.029999999329447746
```

```
  }
}
```

```
        activation: RELU_6
batch_norm {            decay:
0.9997000098228455
center: true            scale: true
          epsilon:         0.0010000000474974513
train: true
        }
       }
      min_depth: 0        max_depth:
0
num_layers_before_predictor: 0
use_dropout: false
dropout_keep_probability:
0.800000011920929
kernel_size: 3        box_code_size: 4
apply_sigmoid_to_scores: false
      use_depthwise: true
     }
   }
   anchor_generator {
ssd_anchor_generator {
num_layers: 6        min_scale:
0.20000000298023224
      max_scale: 0.949999988079071
aspect_ratios: 1.0        aspect_ratios:
2.0        aspect_ratios: 0.5
aspect_ratios: 3.0
      aspect_ratios: 0.3332999944686889 6
     }
   }
   post_processing {
batch_non_max_suppression {
score_threshold: 9.99999993922529e-09
iou_threshold: 0.6000000238418579
max_detections_per_class: 100
      max_total_detections: 100
     }
```

```
  score_converter: SIGMOID
 }
 normalize_loss_by_num_matches: true
 loss                    {
localization_loss {
    weighted_smooth_l1 {
    }
   }
classification_loss {
weighted_sigmoid {
    }
   }
   hard_example_miner {
num_hard_examples: 3000        iou_threshold:
0.9900000095367432        loss_type:
CLASSIFICATION
max_negatives_per_positive: 3
min_negatives_per_image: 3
   }
   classification_weight: 1.0
   localization_weight: 1.0
 }
} } train_config {
batch_size: 16
data_augmentation_options {
 random_horizontal_flip {
 }
}
 data_augmentation_options {
 ssd_random_crop {
 }
} optimizer {
rms_prop_optimizer {
learning_rate {
    exponential_decay_learning_rate                    {
initial_learning_rate: 0.004000000189989805
     decay_steps: 10000
```

```
        decay_factor: 0.949999988079071
       }
     }
     momentum_optimizer_value:
0.8999999761581421      decay: 0.8999999761581421
epsilon: 1.0
   }
 }
  fine_tune_checkpoint:"/home/object-detection-
template/models/ssdlite_mobilenet_v2_coco_2018_05_09/model.ckpt
"   num_steps: 4700   fine_tune_checkpoint_type: "detection"
}
train_input_reader
{   label_map_path:
"/home/Tensorflow/workspace/training_demo/annotations/labelmap.pbtxt"
tf_record_input_reader {
    input_path: "/home/Tensorflow/train.tfrecord"
  } } eval_config {
num_examples: 1
max_evals: 1
  use_moving_averages: false
}
eval_input_reader
{   label_map_path:
"/home/Tensorflow/workspace/training_demo/annotations/labelmap.pbtxt"
shuffle: false   num_readers: 1   tf_record_input_reader {
    input_path: "/home/Tensorflow/eval.tfrecord"
  }
}
```

D. Linguistic variables and fuzzy values

Parameter	Linguistic variables	Fuzzy values
Score	Low	$0<x<30$

	Medium	$25 \leq x < 60$
	High	$55 \leq x \leq 100$
Task difficulty	Easy	$1 \leq x \leq 2$
	Medium	$2 \leq x \leq 3$
	Hard	$3 \leq x \leq 4$
IQ levels	Low	$1 \leq x \leq 70$
	Medium	$71 \leq x \leq 85$
	High	$85 \leq x \leq 200$

E. QUESTIONNAIRE

My name is Rose-Mary Mensah. I am a PhD student at the Kwame Nkrumah University of Science and Technology. The purpose of this questionnaire is to solicit for opinions from teachers and care givers on their perceptions about using social robots as assistive technology to help teachers in the education of children with autism. Your identity throughout this process will be kept strictly confidential. Thank you for participating.

1. Please indicate your gender ○ Male ○ Female

2. Please select your age range

○ 18-30 years ○ 31-40 years ○ 40-50 years

○ Above 50 years

3. What is your highest level of education?

○ Senior High School (SHS) ○ Tertiary (1st Degree)

○ Tertiary(Postgraduate)

4. Do you use technological devices at work when caring for children with autism?

○ Yes ○ No

If yes, please list the gadgets you use

____

5. Have you operated a robot before? ○ Yes ○ No

In some countries such as UK and Germany, research conducted indicates that children with autism are drawn towards technology. As a result, they people have developed robots which are being integrated into classrooms and clinical settings to serve as assistive technology for caregivers. In my research, I wish to gather information on the features caregivers of these special children would like to be incorporated into a robot which is going to be developed for children with autism. Pictures of some of the robots for children with autism are shown below.

6. I would prefer a robot for children with autism to be designed to look like …………………….

○ Human ○

Animal

○ Toy

7. I would prefer the length of the robot to be ……………………….

○ Height of an average adult ○ Half the height of an average adult ○ childlike

8. What level of autonomy would you want the robot to have?

○ Autonomous (robot operates on its own without anyone controlling it) ○ Semi-autonomous (robot can operate on its own and can also be controlled)

○ Wizard-of-Oz (robot would be operated by the caregiver)

1: Strongly Disagree 2: Disagree 3: Undecided 4: Agree 5: Strongly Agree

Performance Expectancy

Item	1 strongly disagree	2 disagree	3 undecided	4 agree	5 strongly agree
1. Robots would be useful in special needs education					
2. Robots would help accomplish teaching tasks more quickly					
3. Robots can increase productivity during school hours					
4. Robots will make it easier for me to do my job					

Effort expectancy

Item	1 strongly disagree	2 disagree	3 undecided	4 agree	5 strongly agree
1. I would find a robot easy to operate					
2. It would be easy for me to learn how to use a robot					
3. Using a robot in my work environment would be easy					
4. My teaching sessions using the robot would be clear and understandable					

Social Influence

Item	1 strongly disagree	2 disagree	3 undecided	4 agree	5 strongly agree
1. My working environment support the use of robots					

2. I would use the robot if people who are important to me think that I should use robot for teaching					
3. I would use robots if people who influence my behaviour think that I should use robots					

Facilitating conditions

Item	1 strongly disagree	2 disagree	3 undecided	4 agree	5 strongly agree
1. I have the required knowledge to operate a robot					
2. Operation costs will encourage the use of robots.					
3. A specific person should be available for assistance with robot operation difficulties					

Behavioural Intention

Item	1 strongly disagree	2 disagree	3 undecided	4 agree	5 strongly agree
I intend to use robots in special needs education in future					

Coefficients table for performance expectancy

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.678	.754		2.226	.031
	ROBOTS WOULD BE USEFUL IN SPECIAL NEEDS EDUCATION	1.097	.454	.858	2.416	.020
	ROBOTS WOULD HELP ACCOMPLISH TEACHING TASKS MORE QUICKLY	-.620	.554	-.462	-1.119	.269
	ROBOTS CAN INCREASE PRODUCTIVITY DURING SCHOOL HOURS	.302	.377	.241	.801	.427
	ROBOTS WOULD MAKE IT EASIER FOR ME TO DO MY JOB	-.157	.430	-.116	-.366	.716

a. Dependent Variable: I INTEND TO USE ROBOTS IN SPECIAL NEEDS EDUCATION IN FUTURE

Coefficients table for effort expectancy

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	1.688	.956		1.767	.084

	I WOULD FIND A ROBOT EASY TO OPERATE	-.314	.329	-.208	-.953	.346
	IT WOULD BE EASY FOR ME TO LEARN HOW TO USE A ROBOT	.695	.321	.505	2.162	.036
1	USING A ROBOT IN MY WORK ENVIRONMENT WOULD BE EASY	-.400	.417	-.257	-.959	.343
	MY TEACHING SESSIONS USING THE ROBOT WOULD BE CLEAR AND UNDERSTANDABLE	.652	.396	.388	1.646	.107

a. Dependent Variable: I INTEND TO USE ROBOTS IN SPECIAL NEEDS EDUCATION IN FUTURE

Coefficients table for social influence

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
	(Constant)	2.161	.560		3.860	.000
	MY WORKING ENVIRONMENT SUPPORTS THE USE OF ROBOTS	-.071	.247	-.066	-.286	.776
1	I WOULD USE ROBOTS IF PEOPLE WHO ARE IMPORTANT TO ME THINK I SHOULD USE ROBOTS FOR TEACHING	.856	.345	.812	2.478	.017
	I WOULD USE ROBOTS IF PEOPLE WHO INFLUENCE MY BEHAVIOUR THINK I SHOULD USE ROBOTS	-.241	.323	-.240	-.745	.460

a. Dependent Variable: I INTEND TO USE ROBOTS IN SPECIAL NEEDS EDUCATION IN FUTURE

Coefficients table for facilitating conditions

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.804	.701		5.423	.000
	I HAVE THE REQUIRED KNOWLEDGE TO OPERATE A ROBOT	.127	.324	.100	.392	.697
	OPERATION COSTS WILL ENCOURAGE THE USE OF ROBOTS	.046	.300	.035	.153	.879
	A SPECIFIC PERSON SHOULD BE AVAILABLE FOR ASSISTANCE WITH ROBOT OPERATION DIFICULTIES	-.003	.199	-.002	-.014	.989

a. Dependent Variable: I INTEND TO USE ROBOTS IN SPECIAL NEEDS EDUCATION IN FUTURE

Rose-Mary Owusuaa Mensah

Department of Computer Science

Kwame Nkrumah University of Science and Technology

Kumasi

25th June, 2018

Dear Parent,

INVITATION TO YOUR WARD TO PARTICIPATE IN RESEARCH

My name is Rose-Mary Mensah, a second year PhD Computer Science student at Kwame Nkrumah University of Science and Technology, Kumasi. I am conducting a study which seeks to research into ways that robots can be used as therapy partners for children with autism. In the research, children will interact with a robot for about 10 minutes per session. All the sessions will be videotaped so that I can gather information about the children's behaviour during interactions with the robot. I would appreciate if you could give permission for your child to participate in the experiments, by signing this consent form. Thank you very much.

Yours Sincerely,

Rose-Mary Mensah

Consent form

I ______________________________, give permission for my child _____________________ to participate in this study.

Date: ______________