INTERACTIVE TOY

Rema MK
Assistant Professor of Computer Science and Engineering Universal Engineering College, Kerala, India
remamk@uec.ac.in

Muhamed Ajmal KR
Department of Computer Science and Engineering Universal Engineering College, Kerala, India
mohdajmal620@gmail.com

Deepak TG
Department of Computer Science and Engineering Universal Engineering College, Kerala, India
dpkofficial3@gmail.com

Roshini M
Department of Computer Science and Engineering Universal Engineering College, Kerala, India
roshinim9296@gmail.com

Muhammed Bazir
Department of Computer Science and Engineering Universal Engineering College, Kerala, India
muhammedbazir7@gmail.com

Abstract*—This is a robotic toy for children to make an interaction with the toy. A smart interactive toy using artificial intelligence which can able to collect data from kids and eventually it will learn everything about the kid like their favorite foods, color etc. It is an AI which is able to make conversation between the kid by using speakers and mic and this improve the communication skill of the kid. This toy also have some camera functions that can give a live tracking to the parents phone and also give a report to parents. It will track the surrounding situations of kid and if any dangerous objects are detected it will alarm and inform the parents. Kids can name the toy according to their will. The toy will also provide awareness on personal hygiene.

Keywords—Message Queuing Telemetry Transport, Real Time Network, Convolutional Neural Network

I. INTRODUCTION

This is a robotic toy for children to make an interaction with the toy. A smart interactive toy using artificial intelligence which can able to collect data from kids and eventually it will learn everything about the kid like their favorite foods, color etc. It is an AI which is able to make conversation between the kid by using speakers and mic and this improve the communication skill of the kid. This toy also have some camera functions that can give a live tracking to the parents phone and also give a report to parents. It will track the surrounding situations of kid and if any dangerous objects are detected it will alarm and inform the parents. Kids can name the toy according to their will. The toy will also provide awareness on personal hygiene.

II. PREVIOUS WORKS

In this paper, introduced a this Barbie contrasts with other dolls that remain motionless and perhaps engage in normal human interactions, including those with children. Barbie is a knowledge hub for educational reasons, which helps youngsters in their schooling and learning and sometimes eliminates the need for any information or teaching support when Barbie is there. These amazing qualities make this interactive Barbie even more captivating. Its intriguing physiognomy includes the ability to start conversations, identify faces and emotions, play solace-inducing music, and convey relevant information to help children feel at ease with their own toys. A safety precaution is to detect dangerous things and barriers near the child or anywhere else to it. In addition, this clicks a warning.[1]

In this paper we demonstrated a service robot navigation system based on the Message Queuing Telemetry Transport (MQTT) protocol communication system that updates the real-time robot states for multiple users. The proposed office assistant robot (OABot) consists of a navigable structure, a mobile app, and a central control workstation and these three components intercommunicate via a wireless network. The voice recognition mobile app is used to interact with users with voice commands; these voice commands are processed inside the workstation and actions are assigned to the moving robot accordingly. The robot can navigate inside the room using real-time maps while localizing itself in the environment. In addition, the robot is equipped with a digital camera to identify people in predefined locations in the room. The WiFi communication system is provided with RESTful and Mosquitto servers for better human-robot communication. Hence, multiple users are notified about the robot status through updates on the real-time states via the MQTT protocol. The developed system successfully navigates to the instructed destinations and identifies the target person with an average accuracy of 96%. Most importantly, in an isolated indoor environment with social distancing restrictions to perform, the proposed system is essentially useful for contactless delivery.[2]

In this paper, the toy business is growing more and more critical of children's privacy issues about smart devices. Parents and guardians continue to work hard to shield their kids from unwarranted privacy hazards such collection, unpermitted usage, and access to their personal data. Unfortunately, in this paradigm, there is still no defined privacy framework that focuses on smart toys, making it challenging to identify potential privacy violations, such as identifying whether or not a phrase communicated with a smart toy is sensitive. We develop a privacy identification method using chatbot technology to address this issue. Children Privacy Identification (CPI) system is the name we give to this system. Our research efforts are divided into two categories to build the CPI system: (1) Take the word from the smart device.[3]

Parent-child separation anxiety frequently occurs when kids start kindergarten and can have a negative impact on both parents' and kids' lives. In parent-child relationships, especially in China, this is a common issue. In order to address the issue, the study I used user research methods to gain a thorough understanding of separation anxiety; ii) created an innovative interaction design that included a physical toy that accompanied children directly and an interactive software that enabled communication between children and parents; and iii) conducted formal user tests to confirm the usability and effectiveness of the proposed design. In particular, BUDDY, a companion product, was created with benefits in the design of an original animal shape, the interaction strategy of communication through time, and the incorporation of fairy stories to create a framework for narrative engagement[4]

Problems with featureless face identification are taken into consideration in this paper. For segmented images, the recognition is based on clustering the closeness measures between brightness cluster cardinality distributions. The Euclidean, cosine, and Kullback-Leibler distances are utilized as proximity measures in this work, respectively. A software model of the recurrent neural network is used to perform proximity measure clustering and image segmentation. The experimental studies of the suggested approach's findings are presented. Currently, content-based image retrieval is extremely pertinent. It is utilized in Internet search engines, biometric identification and technical vision systems, digital image libraries, archives, and databases, among other places. [1 – 3]. The variety of work be addressed in such a integrate face recognition in the search. Face recognition is currently a hot topic, but there is still a long way to go before it is fully addressed. There are three groups that we can separate out of the many different algorithms created for this assignment. In the first of them, identification is carried out by contrasting the distinctive face features [4]. The methods used for this often consist of two steps. The identification and localization of the face in the image are being produced at the initial stage. Face alignment (geometric and brightness), feature extraction, and actually recognition – features matching with the etalons from the data base – are all steps in the second stage procedure. The process of extracting task features is time-consuming and expensive.[5]

Facial recognition has long been a hot topic, particularly now that Covid-19 is so prevalent and less physical touch is required in settings where personnel identification is crucial. Despite the fact that there are numerous facial recognition algorithms, there hasn't been much research into how well these algorithms function when supported by data. The performance of K-Nearest Neighbors (KNN) for face recognition in various scenarios is assessed in this research in order to address this problem. This work attempts to train and test the model using images taken in profile and partially covered faces to imitate the case in which the item required to be identified does not face the camera at a straight angle or wears masks in order to make the results of this study more practical. the experimental findings show with a success probability of 95.0%, K-Nearest Neighbors (KNN) demonstrated greater performance in identifying uncovered frontal faces. However, the model performs less satisfactorily when categorizing profile or masked faces, with a success probability of 22.2% for the former and 2.22% for the latter. It is important to note that the KNN classifier's accuracy for face identification is 100% for frontal faces that are exposed and 74.7% for those that are covered.[6]

Complex decision-making, interpretation, and adaptive planning processes present significant problems for robot intelligence and human-robot interaction (HRI). Recursive task processing and metacognitive reasoning are needed for these. Naturally, the prefrontal cortex, a region of the neocortex, is responsible for realizing these cognitive abilities in the human brain. Prior research on neurocognitive robotics would not satisfy these criteria. The goal is to create a robot control system that can reason spatially, temporally, and emotionally. In this paper, we provide a novel approach that addresses a computer model of the humanoid prefrontal cortex. The focus of computational processes is mostly on the realistic biophysical neuronal structures that are embodied in various dynamics. The dorsolateral, ventrolateral, anterior, and medial prefrontal areas are among the key computational modules that make up the system. It is also in charge of planning the working memory. The working memory sections of the computational prefrontal cortex model are treated using an explainable artificial intelligence (XAI) technique based on reinforcement meta-learning. The created software framework integrated within the humanoid robot platform processes experimental evaluation and verification testing. The created software allows for the observation and control of the humanoid robots' perceptual states and cognitive functions, such as emotion, attention, and intention-based reasoning abilities. To track and assess the performance of the model, several interaction situations are put into practice.[7]

The development of a generic social robotics architecture used in the Nadine social robot, which enables her to convey human-like emotions, personalities, actions, and conversations and can be adapted to suit any scenario or application. Our architecture has three layers: perception, processing, and interaction, and it supports task- or environment-based adaptations as well as modularity (add/remove sub-modules) (for example, change in knowledge database, gestures, emotions). We discovered that it is challenging to conduct a precise state-of-the-art for robots because each one may have been built for a distinct mission or working environment. The possibility of varied hardware on the robots makes comparisons difficult. On the basis of social robot attributes such voice recognition and synthesis, gaze, face, object, affective system, dialogue interaction skills, and memory, we compare Nadine with state-of-the-art robots in this study. [8]

An object-sorting robot system built on the distributed processing framework of the robot operating system. This system can interact with people, perceive a three-dimensional world using a Kinect sensor, reason, and translate natural language commands into machine instructions that control the movement of a manipulator. We specifically suggest a human-robot-environment interactive reasoning technique to boost the robot's intelligence and usefulness. In our approach, the conventional case-based reasoning-belief-desire-intention process is supplemented by a "conversation and 3-D scene interaction module." Our suggested mechanism accomplishes the functions of desire analysis and direction in addition to the conventional function of map matching. Our robot will take the initiative to lead users through discussion when the user's want is lacking or inconsistent with the actual scenario. The user's input data will then be used to refuel the user's desire. The benefits of our mechanism are demonstrated by experimental data. These systems all share the ability of the robot to work alongside people and use a flexible manipulator to do tasks in an unstructured environment. Such a system must be able to perceive its environment, communicate with humans and other robots, recognize targets, use knowledge reasoning, plan its route, seize objects, and perform other tasks. The center of the entire system, the part of an intelligent robot's brain that defines the level of that intelligence, is the reasoning process, which is one of the prerequisites.[9]

Humans have a strong innate capacity for interpersonal interaction. The effectiveness of combined activities depends on the interaction of emergent coordination—a subconscious mechanism based on a close relationship between action execution and perception—and deliberate collaboration. This connection promotes phenomena like reciprocal adaptation, synchronization, and anticipation, which greatly reduce interaction delays and the need for elaborate verbal instructions and lead to the formation of shared goals, which are the basis of social interaction. This is achievable from a neurophysiological point of view because the same brain system that supports action execution also controls interpretation and anticipation of observed action in others. Establishing more natural and effective interaction paradigms with artificial devices, ranging from assistive and rehabilitative technology to companion robots, would require defining which human motion features allow for such emergent coordination with another agent. However, there are significant challenges in examining the neural and behavior mechanisms underlying natural interaction. Particularly, it is highly challenging, if not impossible, for a human agent to limit or regulate quantitatively the unconscious processes that lie at the core of emergent coordination (such as inadvertent motions or looking, for example). However, during an encounter, participants continuously and complexly influence one another, leading to actions that are outside the realm of experimental control. We suggest robotics technol ogy as a viable remedy for this methodological issue in this research. Robots can communicate with a human partner and respond contingently to his behaviors without compromising the experiment's controllability or the naturalness of the interactive situation. To evaluate the sensory and motor mechanisms underlying human human interaction, a robot could serve as a "interactive probe". We illustrate this idea using examples from our work with the humanoid robot iCub, demonstrating how an interactive humanoid robot may be an important instrument for the study of the psychological and neuroscientific underpinnings of social interaction. [10]

The volume of data gathered by the sensory and motor systems is examined in this research. Humans employ their senses to take suitable action after recognizing the information in their environment. Analog electrical impulses are used to transmit information between organs during these procedures. Human sensory organs send information to the brain via analogue electrical signals, and the brain sends commands to the relevant motor organs. This study examines the quantity of analogue signal information produced inside the human body and transforms it into corresponding digital data. In an effort to create a humanoid based on the corresponding digital data, this approach is being used. Based on recent medical articles, the analogue information produced in the human body is examined. These analyses lead to the bit rate and latency specifications for digitally networked nervous systems. This paper demonstrates that when a humanoid performs at a human-like level, both artificial eyes create around 14 Gigabits from a single look. The fact that pressure provides, on average, more information than temperature feeling means that the human body is more sensitive to pressure than to temperature. In order to transform biological signals in the human body into digital information, such as binary bits and bit rate, the amount of information gathered by a human's sensory organs and the commands sent from the cerebrum to the motor organs are analyzed. Analog bio neurological signals are converted to digital values based on reasonable assumptions for the corresponding resolutions and frequencies that humans can distinguish and feel. These analyses can be useful when human like humanoids and artificial limbs are implemented in digital bio-mechanical systems in the future. For example, artificial units that will be used to mimic hum will be converted from analogue to digital values.[11]

In this paper The real-time network (RTNET) is a new network communication approach designed and implemented for humanoid robots that uses five network objects to represent the task and priority of the communication data .the proposed five network objects - alarm, condition, message, mail, and file - are used to represent the task and priority of the communication data. To further enhance the communication mechanism, RTNET can be connected with other protocols like EtherCAT or controller area networks (CAN Bus) for local control systems, such as robot arms. Each subsystem can be connected to the RTNET through Ethernet and communicate with one another there. Additionally, a network structure for the Internet of Things (IoT) based on RTNET is suggested in this study. Users can access all of the IoT network's components, and RTNET is used to gather information from each subsystem. The idea of RTNET is discussed in this study, and RTNET has been used to a humanoid robot control system with CAN Bus at National Taiwan University (NTU).[12]

A state-of-the-art technique to create a machine-learning-based humanoid robot that can work as a production line worker is proposed in the 2017 paper Repeatable Folding Task by Humanoid Robot Worker Using Deep Learning by Pin-Chu Yang, Kazuma Sasaki, Kanata Suzuki, Kei Kase, Shigeki Sugano, and Tetsuya Ogata. The suggested method demonstrates the following qualities: task performing capability, task reiteration ability, generalizability, and ease of applicability. It also offers an intuitive technique to gather data. The suggested method makes use of a head-mounted display to give a first-person perspective and a real-time user interface with a monitor. Teleoperation is utilized through this interface to gather task operating data, particularly for activities that are challenging to complete using a traditional technique. In the suggested method, a two-phase deep learning model is also used. A fully connected deep time delay neural network learns the dynamics of a robot task process using the extracted image features and motion angle data. A deep convolutional autoencoder extracts image features and reconstructs images. The proposed model is tested on an experimental platform with the "Nextage Open" humanoid robot. 35 trained and 5 untrained sensory motor sequences were tested in the object folding task. The trained model performs the item folding task with a success percentage of 77.8% when tested online.[13]

Adaptive Multimodal Emotionndo Detection Architecture for Social Robots was released in 2019 by Juanpablo Heredia, Yudith Cardinale, and Edmu Lopes Silva. A approach for social robots to improve Human-Robot Interaction and model their social behavior is emotion recognition. Since there are numerous ways for people to convey their emotions (such as through their faces, gestures, and voices), multimodal techniques are helpful in supporting the recognition process. Despite work on multimodal emotion identification for social robots, these systems still have limits in the fusion process and perform poorly if one or more modalities are absent or if they have different characteristics. Due of the wide range of sensory abilities of robots, this is a regular occurrence in social robotics; hence, more adaptable multimodal models are required. To help robots better understand human moods in a given environment and adjust their behavior in line with that understanding, we propose an adaptive and flexible emotion recognition architecture that can work with multiple sources and modalities of information and manage varying levels of data quality and missing data. Emotion detection occurs naturally in human social interactions and has a direct impact on how people decide to act and communicate. In social robotics, robots imitate this conduct to communicate with humans in a friendly and natural way. Robots can achieve this through, among other means, recognizing human emotion through visual perception, speech, nonverbal communication and interpersonal contact. For a better grasp of how to interact with people, new concepts for social robots to detect emotions have become more naturalize and quick in recent years. With the expectation of more humanized applications in a more connected world and trends towards the seamless integration of robots into human contexts, deep learning based ideas in particular have increased.[14]

The Adaptive Nonlinear Control Algorithm for a Self-Balancing Robot was introduced in 2019 by yun su, ting wang, and kai zhang. It is a novel composite wheel-leg-track explosive ordnance disposal (EOD) robot with high mobility that can switch between a track and a self-balancing motion mode depending on the environment. In this research, we provide an adaptive nonlinear control approach to enhance the robot's self-balancing mode stability. To construct the nonlinear cascade controller for integrated balance and motion control, we first established a model of the robot's dynamics. Our solution uses a Kalman filtering approach to predict the robot's orientation. Based on this, an adaptive adjustment algorithm adjusts the controller's parameters in real-time in accordance with the robot's condition, improving stability. In order to preserve this stability, we also developed an adaptive zero-offset angle detection technique to account for deviations brought on by changes to the robot's center of gravity (as a result of changes to its mechanical construction). The proposed methods can be used to govern Scorpio in a self-balancing manner, according to the results of tests done to confirm their functionality. We created the Scorpio EOD robot, a revolutionary composite wheel-leg-track device with crawler and self-balancing mobility modes. Scorpio uses crawler tracks in the first mode. Ning Sun served as the associate editor who oversaw the evaluation of this manuscript and gave final approval for publication o ascend stairs and overcome hurdles in challenging environments. When going on flat urban roads, for example, the second mode is chosen because it offers fast speed, flexibility, and low energy consumption requirements. The environmental adaptability and motion flexibility of Scorpio is advantageous, but a more complicated control algorithm is needed to make this hybrid motion work properly. Therefore, the control of Scorpio's self-balancing motion mode—where it essentially functions as a two wheeled self balancing robot (TWSBR)—is the main topic of this study.[15]

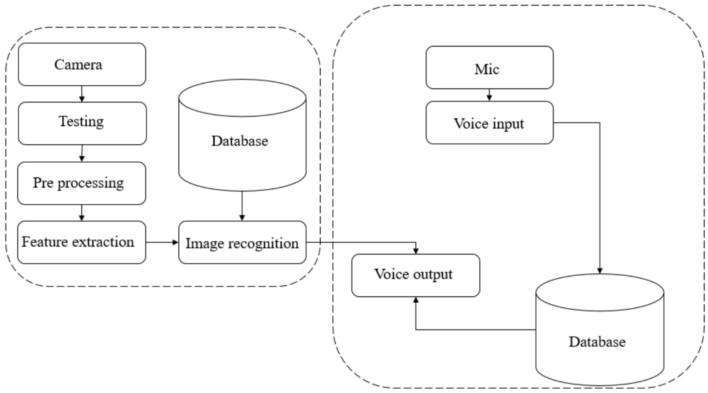
Effect of Robot Embodiment on Satisfaction with Recommendations in Shopping Malls was first published in 2022 by Kazuki Sakai, Yutaka Nakamura, and Yuichiro. Researchers are interested in studying the applications of conversational technology for making recommendations through conversations as a result of recent breakthroughs in this field. In instances where robots discuss objects close to people, it is thought that physical robots, rather than virtual ones, are more effective. Robots may need to suggest elements that are present but invisible in the communication context, though, under real-world circumstances. at this study, we used a field experiment at a mall to look at how robot embodiment affected tasks including recommendations. Following a conversation with participants about their food tastes, the robots made a dish recommendation. We created a conversational recommendation system and put it into use using real-world and artificial intelligence robots. Conversational robots have been created that offer sophisticated information, such recommendations. These machines are capable of offering advanced services while comprehending the emotions and history of the user. Conversational robots are anticipated to be utilized in everyday settings, such as hospitals and as artificial shop clerks Researchers in the area of human-robot interaction have extensively investigated the efficacy of robot embodiment. Studies on the distinctions between real-world and virtual robots have been summarized in recent reviews. According to these surveys, the majority of research found that real robots are more appealing, convincing, and positive than virtual ones, however others found the contrary to be true in specific circumstances. According to one study, virtual robots are more effective for work carried out in virtual environments than physical robots are for duties carried out in real environments. These findings suggest that physical robots are preferable to virtual ones when it comes to promoting products that are actually present where the communication is taking place.[16]

Shih-chang Hsia, Zu_hong Wang, and Bo-yung Wang in 2020 Furthermore, Chuan Yu Chang presented Intelligent Surrounding Recognition for Robot Direction Control In this study, we introduce a brand-new multi sensor-based robot navigation system. The sensors used to build the robot's navigation and monitoring system include a camera, electronic compass, Zig-bee module, and a laser range finder. This solution allows for real-time picture monitoring and remote control of a robot that is on a security patrol in the room. A micro-controller can guide motors to move in the desired directions and at the desired speeds based on information from sensors. The robot can detect obstacles using a 0 to 240 degree field of view laser scanning rangefinder. A wide area can get relative indoor position information from a zig-bee locator. Robot movement direction is controlled by an electronic compass. To enable the robot to move in the middle of the hallway, the computer computes the sensing data and sends the results to the micro-controller. The system has been successfully implemented and tested in a real-world setting. In recent years, indoor environmental safety has become a significant issue in our culture. In the twenty-first century, it is especially clear in wealthy countries that people's lives and property are protected. Limitations and disadvantages exist with the conventional security mechanism. The majority of them work alongside paid security guards. The building's interior spaces may each have a video recorder installed as part of the current security system to keep track of the surroundings. However, they are only able to monitor the fixed area, which has certain blind spots. A smart robot system has recently been designed to carry out environmental safety activities. We can program the robot such that it has enough movement and adaptability to support patrol rounding for a very long time. If there is an emergency, the robot can get here right away and utilize a remote control to provide users crucial videos. One of the key difficulties for a mobile robot in this regard is auto navigation. The pre-schedule or sensing signals decide the robot's movement destination. The robot is guided by the positioned system along the optimum route from its current location to the target place.[17]

Chin S. Chen, Chia J. Lin, and Chun C. Lai published Non-Contact Service Robot Development in Fast-Food Restaurants in 2022. This paper describes the creation of a service robot with cutting-edge improvements in mapping, localization, and navigation that is intended for food service in fast-food restaurants. Additionally, this study took the initiative to create a PC-OGM by combining a 2D occupancy grid map (OGM) with a 3D point cloud map. To put it another way, the service robot can improve its placement and adapt to a more complex environment by using the sensory fusion method. The adaptive motion controller has been improved in terms of navigation so that the service robot can easily go through confined spaces like small aisles. Finally, in order to get input from customers and wait staff, friendly contact-free food service robots were tested at fast food restaurants. For additional study, their comments were divided into three categories: availability, dependability, and satisfaction. We also consider potential future deployment scenarios for food service robots in restaurants to prevent COVID-19 from spreading to food and surfaces. The design and implementation of a food delivery robot, together with its software and hardware system architecture, will be covered in this study. Additionally, this integrated food service robot was created specifically to be used in fast food establishments. The estimation fusion approach is used to incorporate a 3D point cloud map into a 2D occupancy grid map in order to enhance the use of service robots. The creation of PC-OGM inside the combination of the aforementioned two maps aims to increase the robot's spatial orientation and placement accuracy in challenging circumstances. The adaptive motion controller has been improved so that the service robot may more easily traverse through narrow aisles in terms of its navigation function.[18]

III. PROPOSED SYSTEM

Children's artificial intelligence (AI) toy robots are becoming more and more common as interactive and educational toys. They give a variety of activities made to stimulate kids' minds and keep them entertained for a long time. Kids may explore the world of technology, improve their motor skills, and acquire problem solving techniques by interacting with these intelligent robots. robots operate Children's imaginations and cognitive growth are stimulated and engaged by interactive toys. These playthings support a child's creative exploration, learning, and enjoyment. There are many different kinds of interactive toys, including dolls, robots, and even instructional software. Motion sensors, voice recognition, and sound effects are common features of interactive toys. These toys allow kids to communicate with them verbally, physically, or even by tapping a screen. When combined with conventional teaching techniques, interactive toys can improve a child's learning and development. Children can record and keep tabs on actions going on around them with an interactive surveillance gadget. It can be used to record actions, including speaking and moving. The recordings can then be reviewed later or archived for further listening. Interactive surveillance toys are becoming more and more well-liked because they let kids have fun while learning important lessons about digital surveillance. These playthings are excellent for educating kids on the value of confidentiality and security, and they can aid in the development of problem-solving abilities. Robotic surveillance toys are a terrific method to monitor your children while they play. They include a variety of sensors and cameras that let you monitor your children from a distance, keep tabs on their whereabouts, and even set up alarms if they cross predetermined boundaries. Some types even have the capability to communicate with your child or notify you of any emergencies. To keep your children safe and to give you peace of mind, surveillance toy robots are a terrific option.



A. Face Recognition

A camera is used to take a picture of a person's face, which is then processed by the face recognition module of an interactive toy to detect particular facial features like the eyes, nose, and mouth. When the module has identified the facial traits, it can next identify the subject by comparing them to a pre-programmed database of recognized faces. All things considered, a facial recognition module might be a great addition to an interactive toy, enhancing its appeal and personalization for kids.

B. Object Detection

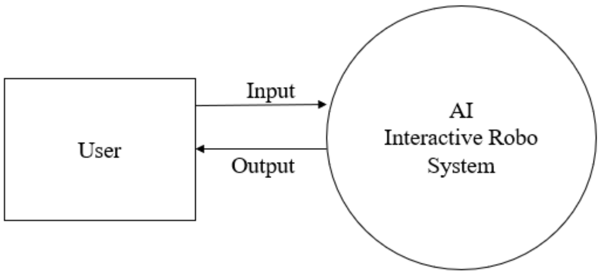
A computer vision approach called object detection involves locating and recognizing particular items within an image or video. Interactive toys are only one of the many uses for this concept. In order to engage with the surroundings and give the user a more engaging experience, interactive toys might employ object detection to identify and react to particular things. A robot that employs object detection, for instance, would be able to identify and communicate with particular toys or objects.

C. Voice Interaction

Voice interaction is becoming increasingly popular for interactive toys, allowing children to engage with their toys in new and exciting ways. Voice interaction refers to the ability of a toy to recognize and respond to spoken commands, allowing children to control the toy through their voice. Voice interaction can provide a more immersive and natural way for children to play with their toys, as they can use their voice to give commands or ask questions, rather than relying on buttons or touchscreens. This can also help to encourage children's creativity and imagination, as they can engage in more open-ended play and storytelling.

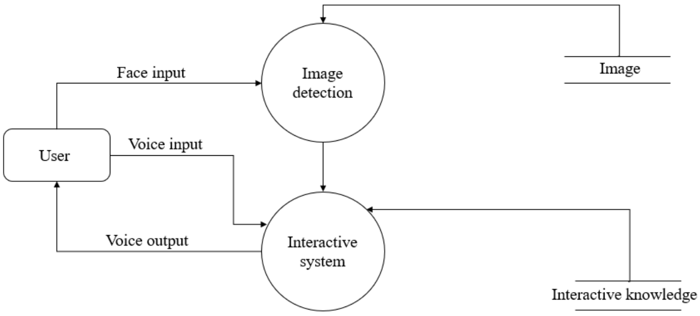
IV. DESIGN

A. *DFD – Level-0*



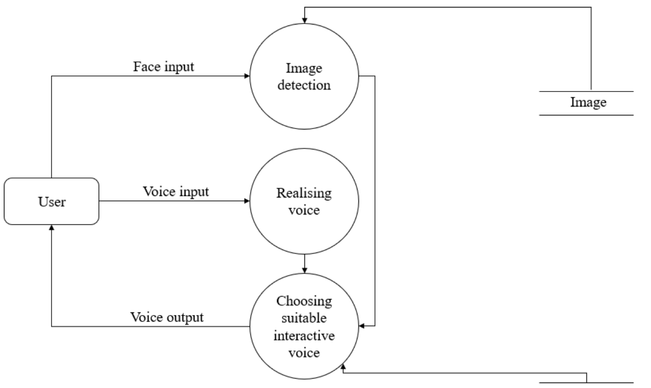
The system would be represented as a single process at DFD level 0, which is the interactive toy. The child's activities, such as speaking commands or touching buttons, as well as any instructions obtained from the toy's creator or seller, could serve as inputs to this procedure. The toy's responses to the child's activities, such as playing music or turning on lights, would represent the system's outputs.

B. *DFD-Level 1*



The user interface, the toy controller, and the toy itself make up the three primary subsystems of a level 1 DFD for an interactive toy. receiving user input and transmitting instructions to the toy controller would fall within the purview of the user interface subsystem. Depending on the unique design of the toy, this can contain buttons, switches, or touch sensors. In order to operate the toy, the toy controller subsystem would be in charge of deciphering user commands and producing control signals. This may entail handling user interface input signals, keeping tabs on the toy's condition, and managing the output signals to the toy.

C. *DFD - Level 2*



A level 2 Data Flow Diagram (DFD) provides a more in-depth look at the processes and data flows that a level 1 DFD discovered. An interactive toy may have a level 2 DFD that deconstructs the functionality of the toy into more detailed steps. The process of how the toy receives and reacts to human input may be shown at DFD level 2. The toy's many sensors that detect user input, the microcontroller that processes the data, and the speaker or LED lights that give the user feedback are all examples of possible data flows.

V. METHEDOLOGY

A. *RESNET*

A deep learning model called Residual Network (ResNet) is applied in computer vision applications. It is an architecture for a convolutional neural network (CNN) that can accommodate hundreds or even thousands of convolutional layers. Previous CNN architectures could only support a few number of layers, which had a negative impact on performance. However, researchers encountered the "vanishing gradient" issue as they added more layers. Gradient descent is used in the backpropagation technique used to train neural networks, which lowers the loss function and identifies the weights that minimise it. The gradient will eventually get so small that it "disappears" if there are too many layers. Performance will also become saturated or worsen with each additional layer. The "skip connections" feature of ResNet offers a novel solution to the vanishing gradient issue. Multiple identity mappings (convolutional layers that initially do nothing; ResNet) are stacked, skipped, and the activations from the preceding layer are reused. By condensing the network into fewer layers, skipping accelerates initial training. The remaining components of the network—referred to as the residual parts—are then free to explore more of the input image's feature space after the network has been retrained, with all layers extended. The majority of ResNet models skip two or three layers at once, with batch normalisation and nonlinearity in between. HighwayNets, a type of more sophisticated ResNet architecture, can learn "skip weights," which dynamically decide how many layers to skip.

B. *YOLOv3*

A real-time object detection system called YOLOv3 (You Only Look Once, Version 3) recognizes particular things in films, live feeds, or still photos. To find an item, the YOLO machine learning system leverages features that a deep convolutional neural network has learned. The YOLO machine learning algorithm has three versions, with the third version being a more accurate version of the first ML method. Versions 1-3 of YOLO were developed by Joseph Redmon and Ali Farhadi. YOLO is a Convolutional Neural Network (CNN) that can quickly identify objects. CNNs are classifier-based systems that can analyze incoming images as organized arrays of data and identify relationships between them (see illustration below). The benefit of YOLO is that it is faster than other networks while still maintaining accuracy. Because it enables the model to view the entire image during testing, its predictions are influenced by the image's overall context. Convolutional neural network methods like YOLO "score" regions according to how closely they resemble predetermined classes. Regions that score highly are reported as positive detections of the class that they most closely match. For instance, using a live traffic feed, YOLO can be utilized to identify various vehicle types. According on which regions of the video score highly in comparison to predetermined classes of cars, YOLO can be utilized, for instance, in a live traffic feed to identify various types of automobiles.

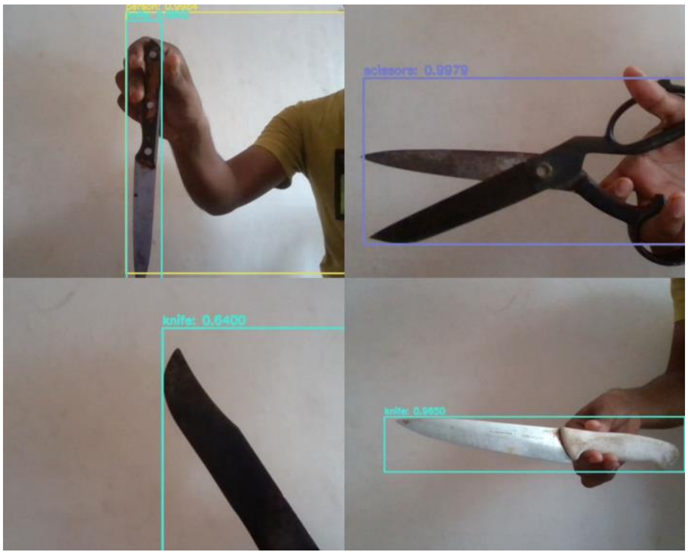
C. *GOOGLE SPEECH RECOGNITION API*

A cloud-based speech recognition service offered by Google is called Google Speech API. It enables programmers to incorporate recognizing speech features into their program, enabling users to communicate with the program using voice commands. Developers can use the Google Cloud Speech-to-Text client library to use the Google Speech API in Python. With the help of this library, users may easily transmit audio files to the Google voice API and instantly receive voice transcriptions in return. Developers must first establish up a project and a Google Cloud account in order to begin. The developer can enable the Google Speech-to-Text API and acquire a set of API credentials after the project is configured. Requests to the API are authenticated using these credentials. The developer can then configure the Google Cloud client library for Python after installing it. In order to send audio data to the Google Speech API and obtain real-time transcriptions, the client library offers a set of functions. By invoking the relevant functions from the client library, the developer can finally include the Google Speech API into their application. The API is capable of handling speech in numerous languages and supports a number of audio formats, including WAV, FLAC, and MP3.Overall, the Google Speech API is a strong tool for giving Python programmers speech recognition capabilities. It is the perfect option for developers wishing to create voice-enabled applications due to its simplicity of use and compatibility for many languages.

VI. FUTURE SCOPE

An interactive toy with modules for face recognition, object detection, and voice interaction has the potential to revolutionize the toy industry by providing a more immersive and interactive experience for children. Here are some potential future scopes for such a toy: Personalized experiences: With the face recognition module, the toy can recognize the child playing with it and personalize the experience accordingly. For example, the toy could greet the child by name, remember their favorite games and activities, and even adjust the level of difficulty based on their past performance. Enhanced interactivity: By incorporating object detection, the toy can detect and respond to different objects in the child's environment. For example, if the child holds up a particular toy or book, the interactive toy could recognize it and respond with appropriate sounds or animations. Natural language processing: With voice interaction, the toy can respond to the child's voice commands and questions, opening up a whole new world of possibilities for interactive play. The toy could also use natural language processing to understand the child's emotions and respond accordingly, providing comfort and support when needed. Educational opportunities: The combination of face recognition, object detection, and voice interaction could create educational opportunities for children. For example, the toy could teach children about different animals by showing them pictures and making the corresponding sounds. It could also quiz children on their knowledge of various subjects, such as math or geography.

VII. DATASET



A well-liked image recognition dataset, the YOLO COCO dataset, is used to train deep learning models, notably for object detection tasks. The dataset known as COCO, or Common Objects in Context, has approximately 330,000 photos that have individually been tagged with details about the objects and their locations. With a single forward pass of a convolutional neural network (CNN), the YOLO (You Only Look Once) method can recognise several objects inside an image and identify their class and position. When training YOLO models for object detection tasks including recognising humans, animals, automobiles, and other typical things in photos, the YOLO COCO dataset is frequently utilised. The COCO Consortium, which is made up of scholars from numerous universities and organisations, is in charge of maintaining the COCO dataset, which was developed by Microsoft. It is widely used in the field of computer vision research and served as the foundation for numerous cutting-edge object identification algorithms.

VIII. CONCLUSION

In recent years, interactive toys have grown in popularity because they give kids fascinating and enjoyable experiences. These playthings are made to encourage kids' imagination, creativity, and cognitive abilities while also encouraging socialization and engagement. Interactive toys have improved in sophistication as a result of technological advancements, giving a variety of functions like voice recognition, artificial intelligence, and internet connectivity. Children can now learn, play, and interact with others in novel and fascinating ways as a result. Interactive toys have already proven to be an effective aid for child development, and as technology advances, we may anticipate seeing these toys get even more sophisticated and cutting-edge in the future. A noteworthy accomplishment in the field of robotics is the creation of an interactive toy robot featuring modules for face recognition, object detection, and voice interaction. Children might learn about technology and develop important abilities like problem-solving, critical thinking, and creativity while playing with such a robot, which would be interactive and informative. The facial recognition module gives the robot the ability to recognize and react to well-known faces, enhancing its capacity for interaction. The robot's capability is further enhanced with the object detection module, which allows it to recognize and engage with items. The vocal interaction module offers an immersive experience by allowing the robot to interact with kids and respond to voice commands. Overall, this kind of interactive toy robot has enormous potential to spark young people's interest in STEM subjects by piqueing their curiosity about robotics and technology. We may anticipate seeing ever more advanced interactive toy robots as technology develops, providing kids with fresh chances to learn and play.

References

[1] Prof. Ravishankar H ,Barbie With Brain: An Interactive Robot https://www.researchgate.net/publication/342497139_BARBIE_WITH_BRAINS_AN_INTERACTIVE_ROBOT

[2] IsankaDiddeniya,IndikaWanniarachchi,HansGunasinghe Huma,Chinthaka Premachandra, Hiroharu Kawanaka human Robot Communication For An Isolated Environment https://ieeexplore.ieee.org/document/9795273

[3] Pei-Chun Lin; Benjamin Yankson; Zhihui Lu; Patrick C.K. HungChildren Privacy Identification System In LINE Chatbot For Smart Toys https://ieeexplore.ieee.org/document/8814570

[4] Yuning Qian; Danni Chang; Yinru Sun BUDDY-A Companion Product Design For Alleviating Parent-Child Separation Anxiety https://ieeexplore.ieee.org/document/9776314
[5] Widodo Budiharto a, Anggita Dian Cahyani b, Pingkan C.B. Rumondor b, Derwin Suhartono aIntelligent Humanoid Robot With Natural Interaction For Education And Entertainment https://doi.org/10.1016/j.procs.2017.10.064
[6] Victor Nemirovskiy,Tomsk Polytechnic University,A.K. Stoyanov,D.S. GoremykinaFace recognition based on the proximity measure clustering January 2016Computer Optics 40(5):740-745 DOI:10.18287/2412-6179-2016-40-5-740-745 https://www.researchgate.net/publication/311695764 _Face_recognition_based_ on_the_proximity_measure_clustering
[7] Ayu Wirdiani , Praba Hridayami, Ayu Widiari, Diva Rismawan,knn Classifier For Face Recognition https://www.researchgate.net/publication/340573794 _Face_Identification_Based _on_KNearest_Neighbor#:~:text=Face%20recognitio n%20utilizes%20facial%20fea tures,an%20instance%2Dbased%20learning%20grou p.
[8] Evren Daglarli ,Computational Modeling of PrefrontalCortex for Meta-Cognition of a Humanoid Robot https://ieeexplore.ieee.org/document/9103513
[9] Manoj Ramanathan, Nidhi Mishra, Nadia Magnenat Thalmann Nadine Humanoid Social Robotics Platform https://www.researchgate.net/publication/333706136 _Nadine_Humanoid_Social _Robotics_Platform
[10] Evren Daglarli ,A Human–Robot-Environment Interactive Reasoning Mechanism for Object Sorting Robot https://ieeexplore.ieee.org/document/9103513
[11] Alessandra Sciutti; Giulio Sandin,Interacting With Robots to Investigate the Bases of Social Interaction https://ieeexplore.ieee.org/document/8068256
[12] Taehyoung Kim; Sungkwon Park,Equivalent Data Information of Sensory and Motor Signals in the Human Body https://ieeexplore.ieee.org/document/9060950
[13] Han-Pang Huang, Jiu-Lou Yan,Tzu-How Huang &Ming-Bao Huang,IoT-based networking for humanoid robots https://doi.org/10.1080/02533839.2017.1372224
[14] Pin-Chu Yang; Kazuma Sasaki; Kanata Suzuki; Kei Kase; Shigeki Sugano; Tetsuya Ogata Repeatable Folding Task by Humanoid Robot Worker Using Deep Learning https://ieeexplore.ieee.org/document/7762066
[15] Adaptive Multimodal Emotionndo Detection Architecture for Social Robots https://ieeexplore.ie,ee.org/document/9705576
[16] Adaptive Nonlinear Control Algorithm for a Self-Balancing Robot https://ieeexplore.ieee.org/document/8945362
[17] Kazuki Sakai; Yutaka Nakamura; Yuichiro Yoshikawa; Hiroshi Ishiguro Effect of Robot Embodiment on Satisfaction With Recommendations in Shopping Malls https://ieeexplore.ieee.org/document/9616371
[18] Shih-Chang Hsia; Szu-Hong Wang; Bo-Yung Wang; Chuan-Yu ChangIntelligent Surrounding Recognition for Robot Direction Control https://ieeexplore.ieee.org/document/9247160