



OPEN ACCESS

EDITED BY

V. R. Sarma Dhulipala,
Anna University, India

REVIEWED BY

Sunita Dhote,
Shri Ramdeobaba College of
Engineering and Management, India
Mohammed Mustaffa,
University of Cape Town, South Africa

*CORRESPONDENCE

Sara Bader Alotaibi
421010101@stu.ut.edu.sa
S. Manimurugan
mmurugan@ut.edu.sa

SPECIALTY SECTION

This article was submitted to
Urban Economics,
a section of the journal
Frontiers in Sustainable Cities

RECEIVED 21 October 2022

ACCEPTED 31 October 2022

PUBLISHED 17 November 2022

CITATION

Alotaibi SB and Manimurugan S (2022)
Humanoid robotic system for social
interaction using deep imitation
learning in a smart city environment.
Front. Sustain. Cities 4:1076101.
doi: 10.3389/frsc.2022.1076101

COPYRIGHT

© 2022 Alotaibi and Manimurugan.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Humanoid robotic system for social interaction using deep imitation learning in a smart city environment

Sara Bader Alotaibi* and S. Manimurugan*

Master of Artificial Intelligence, Faculty of Computers and Information Technology, Industrial Innovation and Robotics Center, University of Tabuk, Tabuk, Saudi Arabia

Introduction: A significant resource for understanding the prospects of smart development is the smart city initiatives created by towns all around the globe. Robots have changed from purely human-serving machines to machines communicating with humans through displays, voice, and signals. The humanoid robots are part of a class of sophisticated social robots. Humanoid robots can share and coexist with people and look similar to humans.

Methods: This paper investigates techniques to uncover proposals for explicitly deploying Artificial Intelligence (AI) and robots in a smart city environment. This paper emphasis on providing a humanoid robotic system for social interaction using the Internet of Robotic Things-based Deep Imitation Learning (IoRT-DIL) in a smart city. In the context of the IoT ecosystem of linked intelligent devices and sensors ubiquitously embedded in everyday contexts, the IoRT standard brings together intelligent mobile robots. IoRT-DIL has been used to create a free mobility mode and a social interaction mode for the robot that can detect when people approach it with inquiries. In direct contact with the actuators and sensors, robotic interface control is responsible for guiding the robot as it navigates its environment and answers questions from the audience.

Results and discussion: For the robots to function safely, they must be monitored and enforced by a central controller using Internet of Robotic Things (IoRT) technology in an emergency. DIL aims to facilitate robot-human interaction by integrating deep learning architectures based on Neural Networks (NN) and reinforced learning methods. DIL focuses on mimicking human learning or expertise presentation to govern robot behavior. The robot's interaction has been tracked in a smart city setting, and its real-time efficiency using DIL is 95%.

KEYWORDS

Human–Robot Interaction, social humanoid robots, smart city, Internet of Robotic Things, deep imitation learning

Introduction

Information and Communication Technologies (ICTs) and technological progress provide new prospects for governing cities more efficiently and comprehensively and for shifting to “smart cities.” Cities increasingly depend on technologies like high-speed internet, 5G networks, IoT, and big data. The introduction of AI and robots in cities has been prompted by the development of smart cities and industrialization. The proliferation of AI technology that can learn through experience—doing more complicated tasks, automating decision-making, and providing assistance in various areas of life—has been made possible by big data and the constantly declining costs of computation and connection (Allam and Dhunny, 2019). Healthcare, transportation (traffic control, advanced driver assistance systems), community security and monitoring systems (image recognition), producing goods (quality assessment), and online shopping are just a few sectors that use AI applications frequently. Together, these factors make AI highly significant in creating smart cities.

Humanoid robots are specialized service robots designed to move and interact like humans. Robots add value in the same manner that all service robots do by automating operations to boost efficiency and reduce overhead. Humanoid robots are an emerging category of robots designed for use in service industries. “Robota” is a Czech word that was used to name robots. Robota directly translates to forced labor, reflecting the general purpose of creating robots to serve man. After all, robots are created to support humans, help them with complex tasks, enhance human activity, and, in some cases, replace humans (Amudhu, 2020). Since robots can aid humans in many fields, this leads to the creation of robots that can support various fields, including drones, military, agricultural, etc.

Since humans constantly need communication, conversations, and social interactions, the general purpose was moving toward creating social robots rather than industrial robots (Smakman et al., 2021), and Socially Assisting Robots SARs were introduced. Social robots are special robots with unique appearances and structures. They have built-in features and qualities that allow them to maximize the human experience, mainly by providing more trust. One of the most apparent features of social robots is their shape since most of them are built in a human-like form. In addition, social robots are taught the social norms and have at least partial automation, making them even closer to humans.

Human–Robot Interaction (HRI) deals with how humankind perceives robots and handles their presence around them (Onnasch and Roesler, 2021). Based on HRI, several robots are employed alongside humans in industrial environments such as car production or assembly factories. On the other hand, a humanoid robot is an advanced form of a social robot. Humanoid robots are unique robots that possess human-like details not only in their physical

appearance, but humanoid robots can show human-like interactions and perform human social skills. These robots, in particular, have software and hardware systems that allow a smoother interaction with humans while using stimulators and manipulators that would enable the free movement of their joints and the capability of understanding human speech. The name “humanoid robot” has gained widespread recognition due to the remarkable resemblance between a humanoid robot and an actual human being. Numerous AI methods, such as NN, DIL, Natural Language Processing, adaptable motor function, and expert systems, are used. The medical, media, academic, and AI research communities might benefit from it.

One of the most commonly used humanoid robots is the Pepper robot. Pepper can analyze people’s voice tone and facial expressions to recognize their respective emotions, allowing the interaction to appear more real (Pandey and Gelin, 2018). In addition, Pepper does not possess sharp edges, making it safe to be around people of all ages and in all sorts of environments. It can interact with its users through a touch screen, speech, light-emitting diodes LED, and tactile head and hands.

The Internet of Robotic Things (IoRT) combines the IoT and robotics to equip robots with cutting-edge capabilities that allow them to thrive in a wide range of settings and applications while opening up exciting new possibilities for businesses and investors. Robots of all kinds can be found in today’s society, from those used in industry and manufacturing to those used in the delivery of food and medical care to those used in law enforcement and even in the cleaning services industry. The transition from having industrial robots working alongside humans in their workplace to social robots interacting with humans within the shared space requires special attention because this interaction has various effects on humans (Lim et al., 2021). This situation makes HRI studies significantly important, especially when discussing how humans perceive robots and how they interact with them. Even in the case of humanoid robots, the machine is still susceptible to errors due to several factors, including recoils and clashes in their joints and having a limited battery or source of power more often than not (Denny et al., 2016).

Since 2008, AI has been incorporated into smart cities research (as a separate topic), according to a scientometric analysis (Ingwersen and Serrano-López, 2018). Additionally, it has been associated with sustainable growth worldwide, notably by developing countries that, for instance, are implementing AI to support the UN sustainability objectives (Apanaviciene et al., 2020). Indeed, supporters claim that AI may be used to address “vicious” urban issues, including pollution, destruction of ecosystems, and quickly rising urban inhabitants (e.g., from industrialization to global warming). AI and robots are projected to permeate every area of human existence and extend across dimensions. They are anticipated, for instance, to play a vital role in “infrastructural development,” to house public assistance, and

to advance the social functioning and wellbeing of smart cities (Golubchikov and Thornbush, 2020).

This research investigates such techniques to uncover proposals for the deliberate integration of artificial intelligence (AI) and robots in smart cities. This study aims to create a humanoid robotic system for social interaction in a smart city utilizing the IoRT and DIL. In the middle of the IoT visualization of linked smart devices and sensors ubiquitously entrenched in everyday contexts, the IoRT standard pulls up intelligent robotic systems.

The rest of the paper is organized as follows: Section Related works on robotics in smart cities describes related research on robotics for smart city environments. Section Proposed humanoid robotic system for social interaction using the IoRT-DIL in a smart city provides an IoRT-DIL for developing a humanoid robotic system with deep learning in a smart city environment. Experimental results and discussion have been given in Section Results and discussion. Finally, the conclusion, limitations, and scope for further research have been shown in Section Conclusion.

Related works on robotics in smart cities

A massive pool of studies on implementing humanoid robots in an intelligent environment has been proposed. In this section, it has been selected a group of similar studies involving humanoid robots, AI, IoT, deep learning techniques, and intelligent cities. The number of papers discussing social robots and their interaction with humans is few, especially in recent years. Even the areas where humanoid robots can be used are many and varied.

In the educational domain, Tanaka et al. (2015) tried to develop a learning application relying on interaction with Pepper and a teacher presenting on its screen. To do that, the developers based their work on care-receiving robot design and total physical response methodology to ensure an interactive experience while learning. The aim was to make it look like Pepper was also learning with the student in the teacher's presence. The application was focused on Japanese children of 5 years old to teach them English from the comfort of their homes. Guggemos et al. (2020) discussed in their paper the acceptance of Pepper robots in the academic field, more specifically as an assistant in higher education classes. The authors aimed to analyze Pepper's adaptiveness, trustworthiness, social appearance, and presence with the classroom students. Lexi was the pepper model used in the study, and the lecturer explained that it could answer the students' questions since the lecturer could not possibly answer 1,500 students.

Qidwai et al. (2020) investigated the benefits of using humanoid robots in assisting teachers in teaching children with Autism Spectral Disorder ASD some academic and social skills.

For this purpose, some children selected by the psychologist and agreed with their parents were chosen to participate, where NAO designed some activities to examine their attention, behavior, and speech. The scheduled programs were also discussed afterward with the teachers to ensure they were suitable for these children.

The Rozanska and Podpora (2019) study discussed the concept of IoT-based humanoid robots and their implementation in human interaction. The authors dealt with the WeegreeOne robot and aimed to advance it to the point that it could be used later in hospitals or assisted living facilities. The robot can perform non-verbal communications with humans based on several qualities, such as speech-text transformation and its ability to interpret human body language and behavior.

To further understand the perception of robots and their acceptance among the standard population, Nyholm et al. (2021) conducted a detailed interview to view the population's perspective on using humanoid robots in healthcare. In healthcare domains, humanoid robots can assist in cognitive training, physical tasks, and monitoring of vital signs. Five female and seven male participants were included in the study to understand their reaction to humanoid robots, whose ages ranged between 24 and 77. The participants were shown a video of Pepper, the humanoid robot, helping in healthcare areas. The video also shows the learning capabilities of Pepper and its design to interact with people.

Arent et al. (2019) proposed an approach to use social robots to diagnose children with autism. The authors here decided to use social interaction between children and the robot combined with a psychological method to assess autism in children. The observations in these experiments must be monitored and evaluated by specialists who are trained psychologists and statistical evaluators upon analyzing recorded videos of the experiments. In their study, Pino et al. (2020) created alternative methods to help patients with mild cognitive impairment fight against cognitive decline. Thus, the effectiveness of social robots such as NAO for these purposes was evaluated.

Podpora et al. (2020) aimed to expand the interaction system of a robot with humans by implementing external smart sensors. The main objective was to gain as much information as possible about humans in sight before an actual engagement between them and the robot takes place, or in other words, to perform interlocutor identification and gaining of parameters. The idea then was to integrate intelligent sensor elements and bring them together through the internet of thing technology to preview a prototype of the system in a small office desk.

Cities have evolved into testing grounds for automation and robot management of urban infrastructure and public areas. The first smart city capabilities are greatly expanded by these Robotic and Autonomous Systems (RAS), as they are termed in engineering. In contrast to so-called "smart" solutions, based on a technical domain in which computers are trained to do tasks,

RAS methods use AI and ML to make judgments and adjust processes to conditions without direct human intervention (Macrorie et al., 2021). An “absence of labor” or “end of jobs” is envisaged as a result of the automated processes in manufacturing as part of smart capitalism, as opposed to AI and robots simply replacing people when it comes to production and managerial choices, as per the work in Peters and Zhao (2019).

Cloud Robotics provided the primary impetus for developing IoRT (Chen et al., 2018). Cloud robotics combines big data technique, computing, and various cutting-edge technologies to build a robot. Cloud robotics facilitates the creation of highly efficient, economic, and power-efficient multi-robot systems that are highly efficient, economical, and power-efficient. A networked robot that can do all tasks, including perceiving, computing, and having a large amount of storage capacity, was produced by Cloud Robotics. This cloud robotics changed from conventional pre-programmed robots to contemporary interconnected robots.

The Internet of Things (IoT), known as IoRT, uses advanced sensors, communication networks, decentralized and localized cognition, and many other components in its most sophisticated form. Robots are intelligent machines in IoRT that are equipped with sensors and can understand what is happening in the outside world by acquiring sensing information from multiple sources and taking the appropriate actions to solve problems. The IoRT has interconnections between elements, abstracted differentiation, scalability, compatibility, autonomous and dynamic robotics, geographic location dispersion, and ubiquitous geographical access (Liu et al., 2020). Field and service robotics are the two primary forms of robotics with which the IoRT apps primarily work (Ray, 2016). Field robots operate in various dynamic and complicated contexts, including forestry, cargo handling, mining, agricultural, and building sites. The field includes the earth's surface, ocean, atmosphere, and space. The service robot helps regular people with domestic chores, workplace tasks, and other tasks. It also helps the aged with medical alternatives. This study claims that the field of Human–Robot Interaction has to take a closer look at how it portrays older people to identify and address any potential bias. The realization that technological progress is a socially constructed process with the potential to reaffirm harmful views of the elderly has led to this acknowledgment and the resulting necessity. Social robots' portrayal of the aging body as “fixable” contributes to a broader ageist and neoliberal narrative in which the elderly are devalued to the role of care receivers, and the burden of providing for their care is transferred from the state to the individual (Burema, 2022).

This study aims to create a humanoid robotic system for social interaction in a smart city utilizing the IoRT and DIL. IoRT offers a platform for various humanoid machines to observe the outside environment, collect sensor data from multiple sources, and use AI approaches to deliver answers by working with the natural world's elements. Robots with

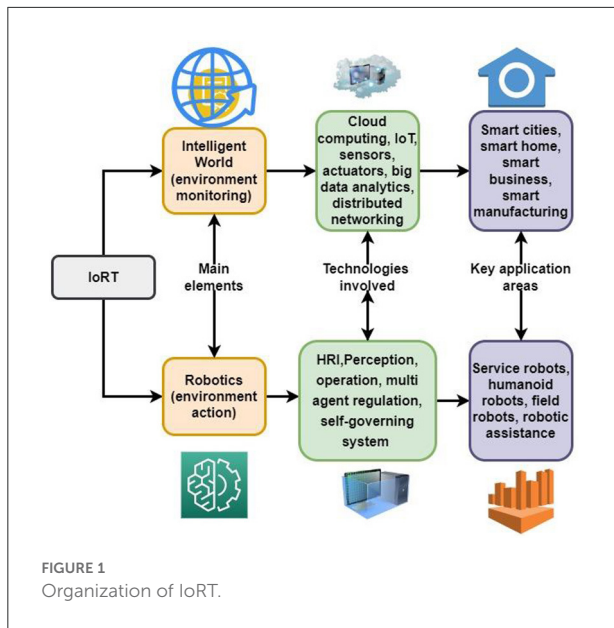
unparalleled degrees of sensing and the capacity to function in unpredictable contexts can use the intuitive DIL method. DIL focuses on mimicking human learning or professional presentation to governing robot interaction.

Proposed humanoid robotic system for social interaction using the IoRT-DIL in a smart city

Humanoid robots are anticipated to be utilized in various contexts, including those with a high potential for danger, such as investigation, search and rescue security, and operations in extreme environments. Improving a few of the robot's most fundamental properties, such as its energy efficiency, will be necessary. The two main categories of robotics that the IoRT deals with are field robots and service robots. Mining, building sites, forestry, cargo handling, and many other fields require field robots because of the complexity and changeability of their environments. It's possible to divide the earth's surface, the ocean, the atmosphere, and outer space into distinct subfields. The main purpose of the service robot is to aid humans in their daily lives by performing tasks such as completing home chores, office work, and healthcare for the elderly. There are four main types of robots used in IoRT applications: those used in the home, those used in the ocean, those used in the air, and those used in the deep sea. The various robot varieties can be studied in depth. IoRT has been constructed on top of the cloud architecture, which uses IoT and cloud computing technologies to give design and implementation freedom for unique robotic applications for interconnected robots that provide services to distant networks. In the IoRT, robots and intelligent environments are combined, as seen in Figure 1.

The intelligent world and robotics components of the IoRT technique are depicted in Figure 1. The foundation of the digital environment is the development of intelligent applications, including those for the smart city, smart home, and smart industries. The primary role of the “intelligent world” is to keep an eye on the region under control using particular sensors created for each area. Energy efficiency is crucial in the intelligent world. One element is lacking in the “intelligent world” despite the reporting capabilities. In other words, it lacks any agents who might walk around and offer services. These agents are known as robots and can be classified as Robotic Associates, Manipulators, Service Robots, or Field Robotic systems, among other uses. A connected service robot with a tray may deliver a customer's food and drink to their table quickly and without hassle. If there is a leak, the tray's sensors will pick it up immediately. The robot is designed to assist wait staff rather than take their jobs.

Modern robotics solely focuses on boosting the robots' awareness level to allow them to carry out their activities



autonomously and comfortably. This increases the automation potential of the robots. Adding behavior controls doesn't usually make the robots more interactive. By utilizing multiple sensors (camera, sonar, and infrared beacons, among others) and computational resources, the IoRT creates a robotic behavioral control. The IoT resources expand the networked robotic functions, enabling the robot to explore and carry out various activities in challenging conditions throughout the intelligent world. The robot's duties change depending on the environment, including navigation, collision avoidance (with static and mobile objects), and encouraging productive interactions between humans and machines. Some of the most significant knowledges in the IoT for the smart world aren't as sophisticated as those in the IoRT. This means the IoRT is more advanced than the IoT.

According to Figure 2, the IoRT architecture comprises five levels. The five stages are the application layer, the robotics layer, the network layer, the internetworking layer, and the physical layer that houses robotic infrastructures. The cloud robotics paradigm is based on the robotics infrastructure layer. In this layer, we find the IoT Cloud Robotics Infrastructure, the Big Data, the IoT Business Cloud Services, and the Robotics Cloud techniques.

Physical layer: The robotic layer comprises hardware components, including robots, sensors, controllers, and smart objects utilized for interaction. In other words, this location has the actual items that may start a conversation. The robots in the vicinity are given a certain amount of perception *via* the sensors. With this layer, the robots can watch the movements and movement of both things and people. The physical layer's actuators carry out routine chores like turning equipment based on the outside weather and other routines.

Network layer: The layer connects the physical layer's components to the internet. Mobile communication networks enabling 3G, 4G/LTE, and 5G are used to connect to the internet. Communication channels like Wi-Fi, Bluetooth, and many others may be used to connect the robots. One can use WiMAX, ZigBee, Z-wave, and Low Power Wide Area Networks (LoRA) to interconnect with robots.

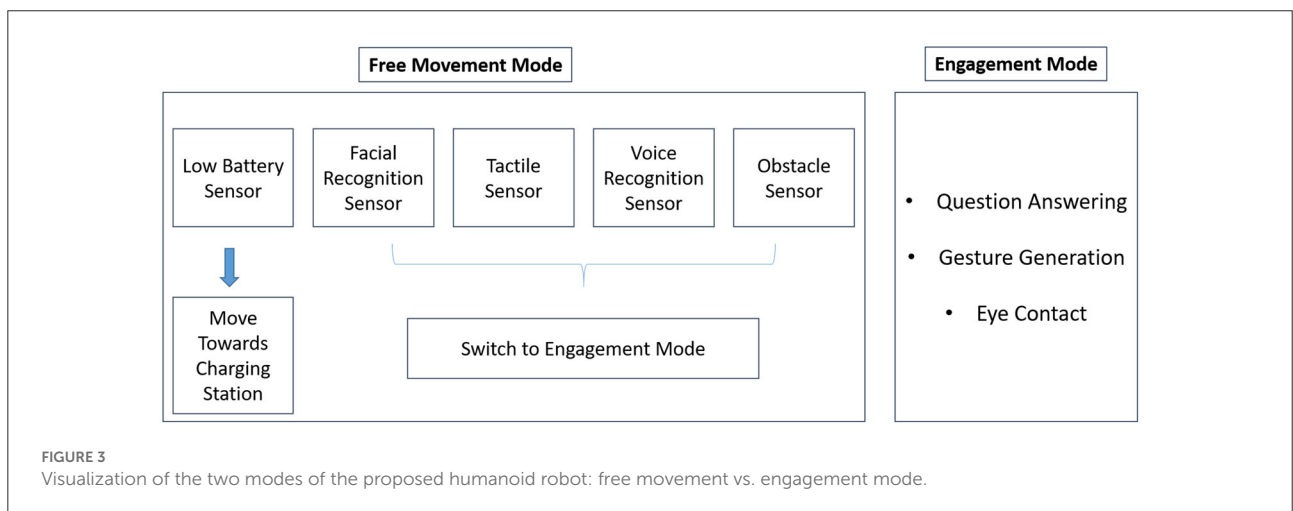
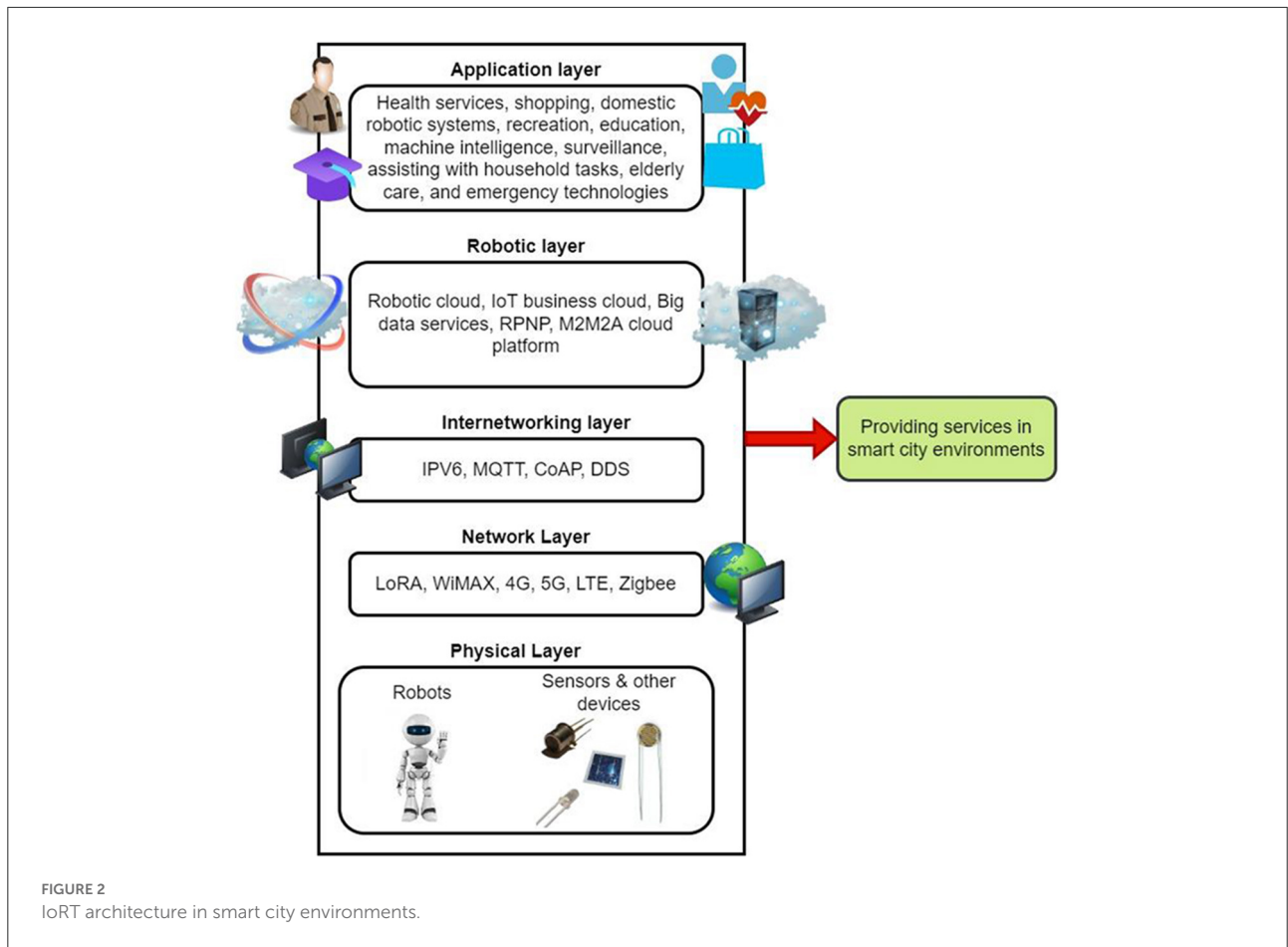
Internetworking layer: The word "IoT" would not occur without the internet. As a result, the internet layer is crucial in developing the IoRT structure. The IoT communication technologies MQTT, CoAP, IPV6, LLAP, and DDS are used to interpret the information between the robots. To give robotic devices internet connectivity, the core network offers to distribute communications, SMS service, and numerous other standards.

Robotic infrastructure layer: The cloud robotics concept is above the robotic infrastructure layer. The Machine-To-Machine-To-Actuator (M2M2A) Cloud Technology, IoT Business, Big Data, and IoT Cloud Robotics are the five main components that make up this layer. M2M2A is primarily made for specialist robots to carry out essential tasks. By fusing robotics and sensor technologies, the M2M2A offers answers to various problems. IoT business cloud services use cloud concepts to solve business-related issues (SaaS, PaaS, and IaaS). Robotics on the IoT infrastructure is used to deliver services.

Application layer: The IoRT architecture's uppermost layer is the application layer. Here, many legitimate solutions have been covered where robots may be used to enhance the customer experience. The humanoid robots can offer various services by engaging in smart city environments like health services, shopping, domestic robotic systems, recreation, education, machine intelligence, surveillance, assisting with household tasks, elderly care, and emergency functions. Future technologies greatly influence the services robots can provide in various applications.

Humanoid robotic system for social interaction using deep imitation learning

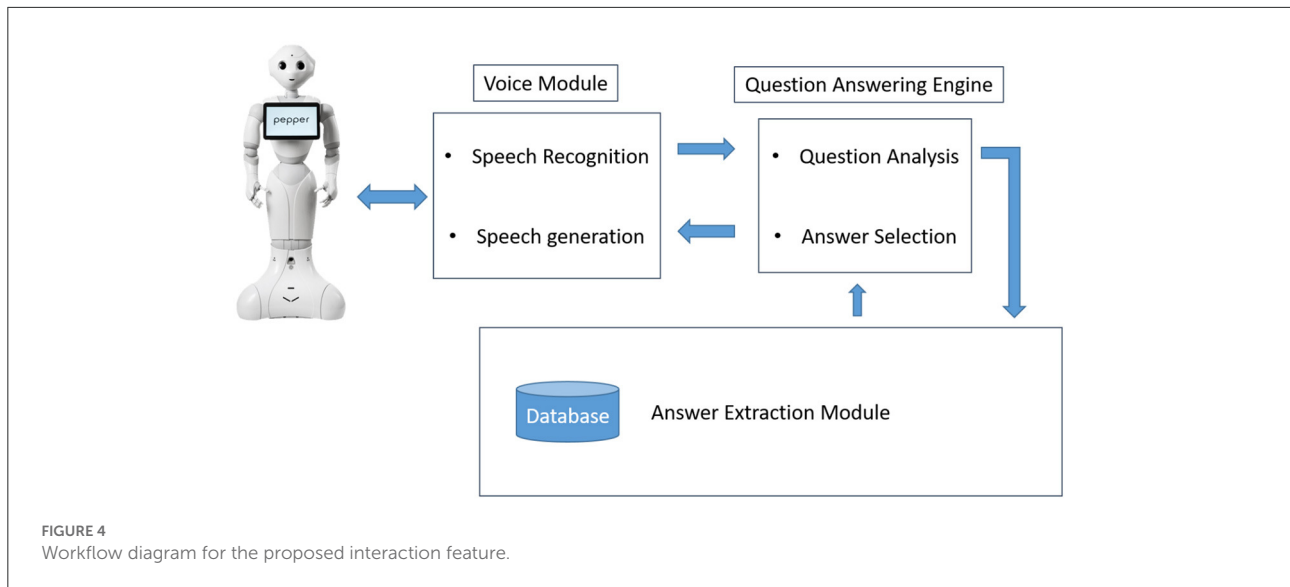
This section discusses the method by which the proposed work will be performed. The proposed system is shown in Figure 3 allows the robot to engage in two modes which are the free movement mode and the social engagement mode. These modes are explained in detail, as well as the robot's interacting ability to answer questions through the voice question/answer model. Imitation Learning (ML) and (DIL) to determine how robots behave. DIL combines the ideas of deep learning architecture and NN with deep learning algorithms. DIL aims to regulate robot behavior in ways that mimic human learning and expert demonstration.



Free movement mode

The first step involves the increased engagement between robots and humans, which can be represented in the robot’s movement in the smart city, gestures, eye contact, and speech capabilities. Initially, the robot will be roaming the city with its

senses activated. The robot is equipped with voice detection, facial recognition, and touch sensors, which means that when the robot is touched, the system will convert into engagement mode, meaning that a human might need to interact with the robot. Similarly, the robot would gain knowledge when a human



is moving toward it or if a human is speaking to it. Furthermore, the robot will be able to not bump into people or obstacles around it through instilled cameras. Additionally, the robot can move to its charging station when it senses a low battery.

Engagement mode

The engagement mode is activated when a human presence is detected. In this case, the robot introduces itself as an assisting robot that can answer and interact with people in a smart city environment. The human can interact with the robot through voice, and it can also reply with voice.

Interaction mode

The proposed system shown in [Figure 4](#) comprises four main modules: the Speech recognition/generation module, the question-answering engine, the answer module, and the update module.

Initially, the robot receives a question through human speech recognition. Next, the robot relies on a question analysis technique to extract the keywords used in the question and relate them to the answer database. To do so, the developers have developed an answer database containing general information about the smart city environment. After analyzing the question, an answer is developed and chosen, which is communicated with the Q/A engine and then with the speech module to generate a speech representing the spoken answer. Additionally, suppose the robot faces a new vocabulary it doesn't understand. In that case, it will activate the touch screen asking the human to add this new word to the database for the update. The interaction of the proposed humanoid robot in a smart city environment has been shown in [Algorithm 1](#):

```

Step 1: Welcome message
Step 2: Get user input
Step 3: If user asks questions,
    Text processing using BNF
    If available in the knowledge database,
    Speak "the answer" to the user and go to step 2
    End
Step 4: If the answer is not available in the knowledge database
    speak "no answer found"
    Get the answer and store it in the knowledge database
    Else go to step 2
    End
End

```

Algorithm 1. The interaction of the proposed humanoid robot.

It is also essential within the engagement mode to maintain eye contact with the human and perform body gestures to make the humanoid robot more engaging and trustworthy. To develop the question-answering model, a group of sentences relating to the topic (smart city environment) must be collected. These sentences are in the form of questions and answers. Next, sentence patterns are defined. Penalties are defined in a formal query or punishment based on the samples above. The practice, in this case, is similar to fill-in-the-blank. This means that the pattern has some variable slots that can be translated into class identifiers or interrogatives.

After that, classes are assigned, and Backus Normal Form (BNF) grammar is built, which involves mapping the sentence pattern into standard BNF form. Then, the BNF grammar is translated to word strings through the BNF parser. When the robot receives a question, part of speech tagging, answer type tagging, and parsing occur as processing steps.

Deep imitation learning

Deep Reinforcement Learning (DRL) is an approach the robot uses to attain its objectives incrementally. For instance, an automatic robot vehicle’s movements must be optimized to reach the target sequentially. The system penalizes the robot for poor performance and compensates it for good performance. For some tasks, reinforcement learning requires more time to grasp the results of each activity performed by the robot. DIL uses a Deep Q-Learning method to circumvent the time delay caused by DRL.

As the name suggests, imitation involves observing how others behave and replicating them in a live setting. For example, the humanoid robot mimics a human grabbing a milk carton before actually doing it. Consequently, DIL may be seen as an ability that involves learning new skills by imitating living beings or other machines. DIL aids the humanoid robot by incorporating human understanding into it and assisting it in processing the data collected by the sensors to build a conceptual model. This practice is sometimes referred to as understanding through illustrations. The robot will receive a positive incentive if it operates well in the outside world; otherwise, it will receive a lousy incentive/penalty. The robot uses DIL to determine which activity results in an insufficient incentive. By pretraining the robot using collective knowledge input, DIL enables the robot to function well in challenging circumstances. If the automated robotic vehicle is not taught correctly, it would likely cause fatal crashes.

Markov process has been used to model the problems in DIL, and for any movement it takes, it has to gratify the equation as follows:

$$P\{f_{m+1}|f_1, b_1, \dots, f_m, b_m\} = P\{f_{m+1}|f_m, b_m\} \quad (1)$$

Equation (1) shows the f is the learning state, b is the action, and $P(f_1)$ is the initial probability distribution function. The incentive function for the robot using DIL can be represented in Equation (2),

$$I(f_p, b_p) : F * B \rightarrow I \quad (2)$$

The term “Action” may be characterized as a list of options from which the robot can select, such as moving ahead, turning around, answering, etc. The symbol “ α ” denotes the discount rate and the scale of the potential benefits. The existing incentive is represented by $\alpha = 0$. The protracted incentive is represented by $\alpha = 1$. The environment (smart city) is the setting where the robot travels, using the agent’s present behavior and the state as input. The location where the robot is engaged during the progressive movement may be used to illustrate the state. The incentive comes from the environment’s input, based on the winning and losing rates of each activity the robot performed. The regulation, represented as β , is a probabilistic function that

aids in choosing the following course of action depending on the present state to ensure a larger incentive.

$$\beta_\tau : F \rightarrow P(B) \quad (3)$$

Equation (3) displays $P(B)$ as the likelihood of action B . $\beta \in I^p$ defines the regulation $\beta_\tau(f_p, b_p)$. The regulation creates a value for the robot to identify the realistic environment and produces the path having state, incentive, and action as follows in Equation (4):

$$\varepsilon_{1:p} = \{F_1, b_1, I_1, \dots, F_m, b_m, I_m\} \quad (4)$$

The performance of the robot has been evaluated using the equation given below:

$$I_p \alpha = \sum_{l=m}^{\infty} \alpha^{l-m} I(f_l, b_l) \quad (5)$$

Equation (5) shows that α is the incentive factor, and its value lies between 0 and 1. I_p is the incentive at time p . $I(f_l, b_l)$ is the incentive function obtained during the state f_l and action b_l . Since DIL involves Q-learning, the Q value maps the state and action with the incentive function based on the regulation function β . It is represented as

$$Q_{fun} = Q_\beta(F, B) \quad (6)$$

Equation (6) expresses the optimal path for the robot to travel to produce its output chosen by the created Q value. The state-action pairing is mapped to the incentives using a NN technique in DIL, which then modifies the weights and variables to provide an error-free result. The regulation of DIL is to map the best course of action depending on the present state to ensure a larger incentive and is represented as:

$$B = \beta(F) \quad (7a)$$

$$Q_\beta(F_p, B_p) \leftarrow (1 - \delta) * Q(F_p, B_p) + \delta * [I_p + \alpha * \max Q(F_{p+1}, B_{p+1})] \quad (7b)$$

Equations (7a) and (7b) illustrations F_p is the learning state, B_p is the action at the current state p . (F_{p+1} and B_{p+1} are the state and action during the next state $p + 1$. δ is the learning rate of the proposed humanoid robot. β is the regulation for choosing the following course of action depending on the present state to ensure a more considerable incentive. The symbol “ α ” denotes the discount rate; the potential incentives scale up.

A set of input-output functions (n, s) that illustrate the quantitative measure of DIL show the state of the demonstration (smart city) environment in n and the demonstrator’s response with s . A form of tuple expresses the robotic behavior (state, action, incentive, next state). The presentation $Pres = [n_i, s_i]$, where the robot learns the regulation β , which is denoted by

$$Pres(p) = Pres[y(p), q, \beta] \quad (8)$$

Equation (9) shows the p is the amount of time needed to complete the next activity, y is the feature extraction method, q is the anticipated motor activity, and β is the set of regulations employed. DIL significantly shortens the robot's learning curve. The DIL enables the robot to accomplish tasks autonomously while learning from its expertise.

Results and discussion

By observing how a service robot interacts with humans and various barriers in a smart city setting, an experimental arrangement is made to track the robot's activities. The Arduino MEGA2560 board is used to execute the robot, and the Arduino Software is used to program it (IDE). Due to the changing surroundings, quickly moving people and objects, and other factors, the robot finds navigating to be a little bit challenging. Using a motion controller and optical navigation, the robot can understand its surroundings without colliding with people, moving items, or stationary things. The goal of the smart city is to install robots in public spaces like dining establishments, healthcare, and shopping centers to interact with the general population and provide services. The intelligent robots communicate with the city's residents, promising them a good life. The robots' IoRT technology employs a microcontroller to supervise them in case of emergencies and to carry out any enforcement necessary to prevent mishaps. An experimental setup is developed to track the robot's actions as it interacts with clients and other impediments in a normal restaurant.

The robot displays these activities in the smart city environment, recognizing its states and responding to input by firing an action. The robot's behavior alters as each state does. If someone gestures at the robot, it should adjust its actions to approach them and follow its instructions. The robot should approach people when it sees them and provides services. People in smart city environments can engage with the robot by scanning a QR code or using speed. Comparison has been made with the human-human interaction and the proposed robotic system interaction using IoRT-DIL in terms of the reply time (during interaction), speed of movement (mobility), and performance efficiency (based on algorithms used).

Figure 5 shows the comparison of reply time among humans and the proposed robot using IoRT-DIL during interaction for various experiments. In trial 1, the proposed robot had a poor or increased reply time than the human reply time for all the experiments. However, during trial 2, the reply time of the proposed robot improved and was close to the human response time. Also, initially for a few experiments, humans and robots have reduced reply time, after which the reply time increases with the number of experiments done. In general, the reply time of the robot in trial 2 is better than the human reply time.

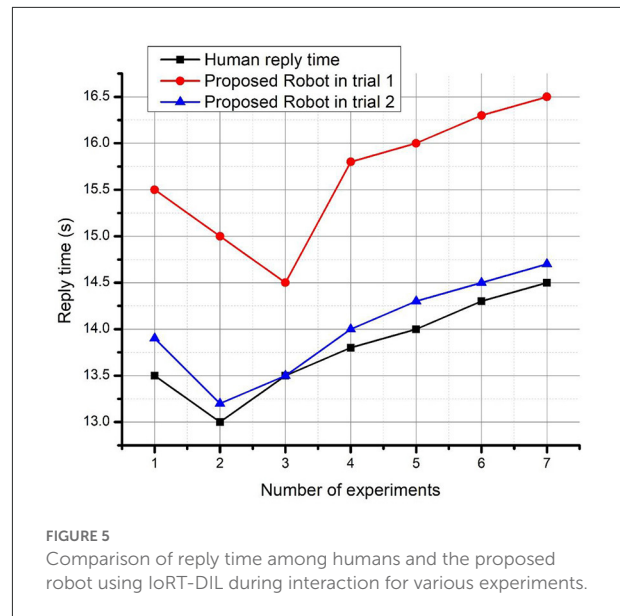


FIGURE 5 Comparison of reply time among humans and the proposed robot using IoRT-DIL during interaction for various experiments.

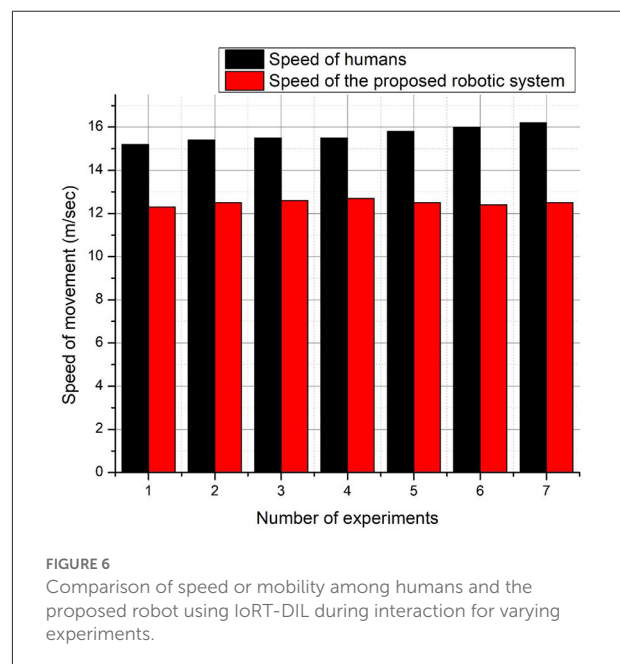


FIGURE 6 Comparison of speed or mobility among humans and the proposed robot using IoRT-DIL during interaction for varying experiments.

Figure 6 depicts the comparison of speed or mobility among humans and the proposed robot using IoRT-DIL during interaction for the varying number of experiments. The mobility of the proposed robotic system using IoRT-DIL is limited due to the delay in acquiring the signals from various sensors or devices and processing them during the movement mode of the robot. Also, the robot's mobility remains almost the same for all the experiments. However, mobility is high for humans and has increased with the number of experiments.

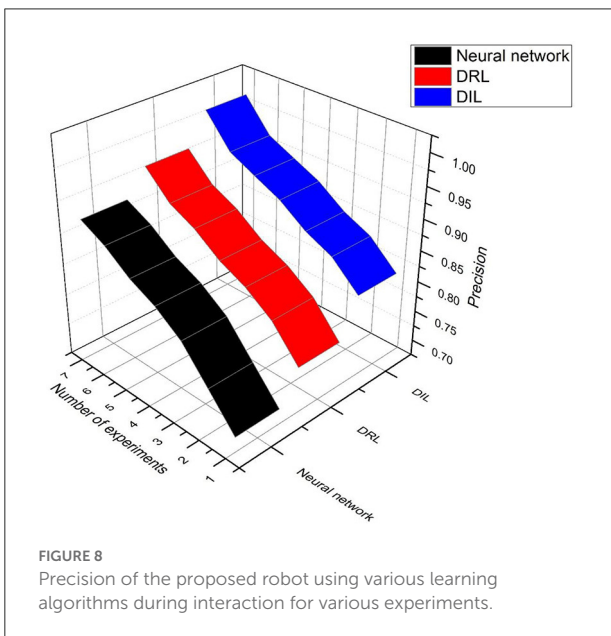
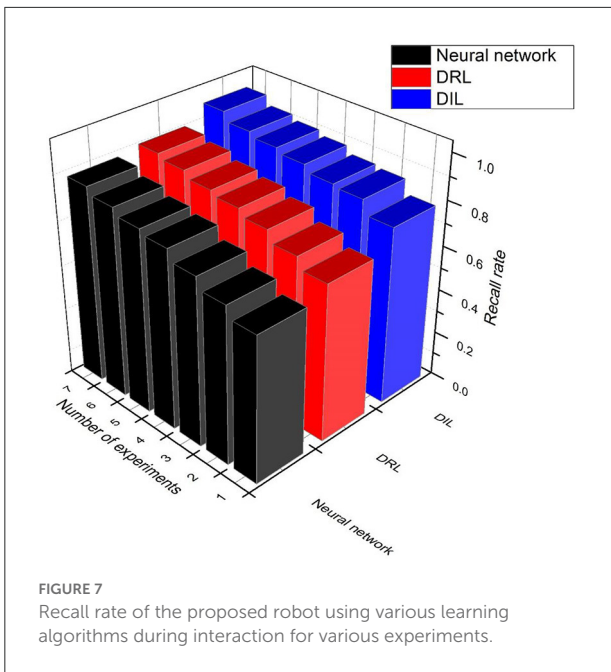
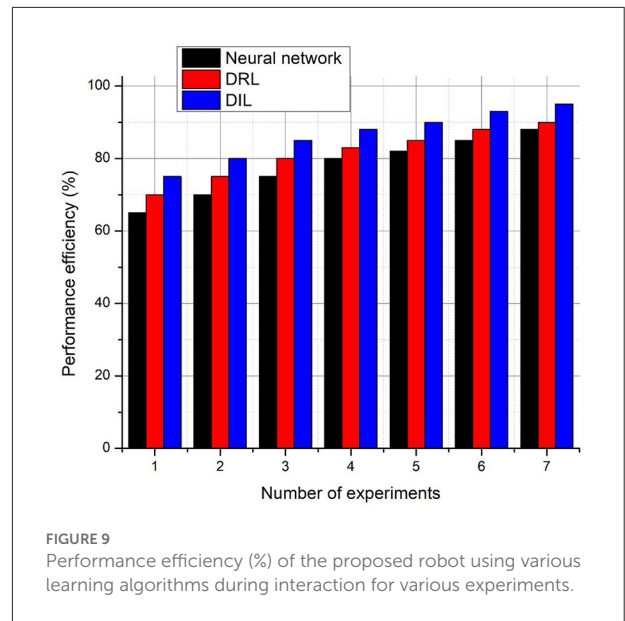


Figure 7 shows the recall rate of the proposed robot using various learning algorithms during interaction for various experiments. The recall rate has been used to analyze the performance of the various learning algorithms in the proposed robotic system. The recall rate has been computed as follows:

$$Recall = \frac{\sum CI}{\sum AT} \tag{9}$$

Where *CI* refers to the correct interaction of the proposed robot by answering the queries effectively in a smart city



environment. *AT* refers to the actual truth that exists in a smart city environment. It has been observed that as the number of experiments increases, the recall rate improves as the knowledge base improves. The neural network has a reduced recall rate among the different learning algorithms, whereas the proposed DIL has the highest recall value of 0.97 for experiment 7.

Figure 8 shows the precision of the proposed robot using various learning algorithms during interaction for various experiments. Precision has been used to analyze the performance of the various learning algorithms in the proposed robotic system and has been computed as follows:

$$Precision = \frac{\sum CI}{\sum TNI} \tag{10}$$

Where *CI* refers to the correct interaction of the proposed robot by answering the queries effectively in a smart city environment. *TNI* refers to the total number of interactions done in a smart city environment. It has been observed that as the number of experiments increases, the precision improves as the robot gets better training and learning experience. Among the different learning algorithms, the neural network has low precision, and DRL has moderate performance in terms of precision. However, the proposed DIL achieved the highest precision value of 0.99 during experiment 7.

Figure 9 depicts the proposed robot's performance efficiency (%) using various learning algorithms during interaction for varying experiments. The deep learning algorithm using a neural network gave poor efficiency for all the experiments. The DRL algorithm gave improved performance than deep learning with a peak efficiency of 91% for experiment 7. This enhanced performance of DRL has been attributed to its optimization to reach the target sequentially. The proposed DIL has the

highest efficiency (with a value of 95%) than DRL. The reason is attributed to using a Deep Q-Learning method in DIL to circumvent the time delay caused by DRL.

Conclusion

The IoRT is a relatively new concept in information and communication technology that aims to improve people's daily lives by applying cutting-edge technological advancements in robotics and IoT. As a robust method for handling robot control, robotic behavioral control for IoRT is a game-changer. Robotic behavior is programmed with deep learning as the guiding principle. For robots to achieve previously unimaginable levels of observation and adaptability in challenging settings, deep learning is a natural method of instruction. A humanoid robotic system for social interaction by means of the IoRT-DIL in a smart city has been developed. IoRT-DIL has been used to create a free mobility mode and a social interaction mode for the robot that can detect when people approach it with inquiries. The tasks required to move around the surroundings and respond to queries from the audience are controlled by robotic interaction control; this links the sensing and acting components directly. To facilitate robot-human interaction, DIL combines the concepts of deep learning architecture, wherein NN are used, and reinforcement learning methods. DIL focuses on mimicking human learning or expertise presentation to govern robot behavior. It will take a long time to teach the robot if the compensation signal it receives is incorrect. Using the lens of behavior and cognitive sciences, IL teaches robots to adapt their actions in response to their surroundings. The robot is placed in a simulated restaurant setting where its various behaviors are observed through customer interactions. Comparison has been made with the human-human interaction and the proposed robotic system interaction using IoRT-DIL in

terms of the reply time (during interaction), speed of movement (mobility), and performance efficiency (based on algorithms used). The proposed robot's interaction using DIL has been tracked in a smart city setting, and its real-time efficiency is 95%.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SA: conception and design of study. SM: acquisition of data and analysis and interpretation of data. Both authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Allam, Z., and Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. *Cities* 89, 80–91. doi: 10.1016/j.cities.2019.01.032
- Amudhu, L. T. (2020). "A review on the use of socially assistive robots in education and elderly care," in *Materials Today: Proceedings*.
- Apanaviciene, R., Vanagas, A., and Fokaides, P. A. (2020). Smart building integration into a smart city (SBISC): development of a new evaluation framework. *Energies* 13, 2190. doi: 10.3390/en13092190
- Arent, K., Kruk-Lasocka, J., Niemiec, T., and Szczepanowski, R. (2019). "Social robot in diagnosis of autism among preschool children," in *2019 24th International Conference on Methods and Models in Automation and Robotics (MMAR)*. IEEE, 652–656.
- Burema, D. (2022). A critical analysis of the representations of older adults in the field of human-robot interaction. *AI Society* 37, 455–465. doi: 10.1007/s00146-021-01205-0
- Chen, W., Yaguchi, Y., Naruse, K., Watanobe, Y., Nakamura, K., and Ogawa, J. (2018). A study of robotic cooperation in cloud robotics: architecture and challenges. *IEEE Access* 6, 36662–36682. doi: 10.1109/ACCESS.2018.2852295
- Denny, J., Elyas, M., D'Costa, S. A., and D'Souza, R. D. (2016). Humanoid robots—past, present and the future. *Eur. J. Adv. Eng. Technol.* 3, 8–15. Available online at: <https://ejaet.com/PDF/3-5/EJAET-3-5-8-15.pdf>
- Golubchikov, O., and Thornbush, M. (2020). Artificial intelligence and robotics in smart city strategies and planned smart development. *Smart Cities* 3, 1133–1144. doi: 10.3390/smartcities3040056
- Guggemos, J., Seufert, S., and Sonderegger, S. (2020). Humanoid robots in higher education: evaluating the acceptance of pepper in the context of an academic writing course using the UTAUT. *Br. J. Educ. Technol.* 51, 1864–1883. doi: 10.1111/bjet.13006
- Ingwersen, P., and Serrano-López, A. E. (2018). Smart city research 1990–2016. *Scientometrics* 117, 1205–1236. doi: 10.1007/s11192-018-2901-9
- Lim, V., Rooksby, M., and Cross, E. S. (2021). Social robots on a global stage: establishing a role for culture during human-robot interaction. *Int. J. Soc. Robot.* 13, 1307–1333. doi: 10.1007/s12369-020-00710-4
- Liu, Y., Zhang, W., Pan, S., Li, Y., and Chen, Y. (2020). Analyzing the robotic behavior in a smart city with deep reinforcement and imitation learning using IoRT. *Computer Commun.* 150, 346–356. doi: 10.1016/j.comcom.2019.11.031

- Macrorie, R., Marvin, S., and While, A. (2021). Robotics and automation in the city: a research agenda. *Urban Geogr.* 42, 197–217. doi: 10.1080/02723638.2019.1698868
- Nyholm, L., Santamäki-Fischer, R., and Fagerström, L. (2021). Users' ambivalent sense of security with humanoid robots in healthcare. *Inform. Health Soc. Care* 46, 218–226. doi: 10.1080/17538157.2021.1883027
- Onnasch, L., and Roesler, E. (2021). A taxonomy to structure and analyze human–robot interaction. *Int. J. Soc. Robot.* 13, 833–849. doi: 10.1007/s12369-020-00666-5
- Pandey, A. K., and Gelin, R. (2018). A mass-produced sociable humanoid robot: pepper: the first machine of its kind. *IEEE Robot. Automat. Magazine* 25, 40–48. doi: 10.1109/MRA.2018.2833157
- Peters, M. A., and Zhao, W. (2019). “Intelligent capitalism’ and the disappearance of labour: whitherto education?,” in *Education and Technological Unemployment* eds Peters, M. A., Jandric, P., Means, A. J. (Singapore: Springer), 15–28. Available online at: <https://link.springer.com/content/pdf/bfm:978-981-13-6225-5/1.pdf>
- Pino, O., Palestra, G., Trevino, R., and De Carolis, B. (2020). The humanoid robot NAO as trainer in a memory program for elderly people with mild cognitive impairment. *Int. J. Soc. Robot.* 12, 21–33. doi: 10.1007/s12369-019-00533-y
- Podpora, M., Gardecki, A., Beniak, R., Klin, B., Vicario, J. L., and Kawala-Sterniuk, A. (2020). Human interaction smart subsystem—extending speech-based human-robot interaction systems with an implementation of external smart sensors. *Sensors* 20, 2376. doi: 10.3390/s20082376
- Qidwai, U., Kashem, S. B. A., and Conor, O. (2020). Humanoid robot as a teacher's assistant: helping children with autism to learn social and academic skills. *J. Intelligent Robot. Syst.* 98, 759–770. doi: 10.1007/s10846-019-01075-1
- Ray, P. P. (2016). Internet of robotic things: concept, technologies, and challenges. *IEEE Access* 4, 9489–9500. doi: 10.1109/ACCESS.2017.2647747
- Rozanska, A., and Podpora, M. (2019). Multimodal sentiment analysis applied to interaction between patients and a humanoid robot Pepper. *IFAC PapersOnLine* 52, 411–414. doi: 10.1016/j.ifacol.2019.12.696
- Smakman, M. H., Konijn, E. A., Vogt, P., and Pankowska, P. (2021). Attitudes towards social robots in education: enthusiast, practical, troubled, sceptic, and mindfully positive. *Robotics* 10, 24. doi: 10.3390/robotics10010024
- Tanaka, F., Isshiki, K., Takahashi, F., Uekusa, M., Sei, R., and Hayashi, K. (2015). “Pepper learns together with children: development of an educational application,” in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, Vol. 2015, 03–05. doi: 10.1109/HUMANOIDS.2015.7363546