DOI: 10.15849/IJASCA.241130.11

Int. J. Advance Soft Compu. Appl, Vol. 16, No. 3, November2024
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Human-Robot Interaction (HRI) using Machine Learning (ML): a Survey and Taxonomy

Mohammad Adawy[1], Hasan Abualese[2],
Nidhal Kamel Taha El-Omari[3], Abdulwadood Alawadhi[4]

[1, 3]Faculty of Information Technology
The World Islamic Sciences and Education University (WISE), Amman, Jordan
[2] Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan.
[4]School of Computing, Universiti Utara Malaysia, Sintok, Malaysia
[1]Mohammad.adawi@wise.edu.jo, [2] h.abualese@ammanu.edu.jo,
[3]Nidhal.omari@wise.edu.jo, [4]alawadhi@uum.edu.my

Abstract

Human-robot interaction (HRI) which has become the fundamental need of the hour is born out of the necessity for studying the relation between humans and robots. This cutting-edge discipline is a multidisciplinary field that draws from computer science, robotics along with human-computer interaction and psychology. It focuses mainly on designing and programming machines, best known as automated machines or robots, which are used by humans to perform specific tasks in a timely manner and with higher quality. The key problem in HRI is to realize, shape, tune, and modelling the human-robot interaction in a flexible manner. For the sake of reflecting and shaping the interactions between humans and robots, HRI is based on the fusion of the two areas: the people's behaviour and attitudes towards using these robots, as well as the physical, technological, and interactive features of the robots. As the robot has tightly integrated from a set of sensors that collect the data from the environment and send them to the processor which in turn translates the collected data into information that can be used in the robot itself, machine learning (ML) is a well-known research area that focuses on the building of well-stocked knowledge systems by using supervised and unsupervised algorithms. From a conceptual standpoint, this research survey and taxonomy pursue to present an in-depth evaluation and review of the most current state-of-the-art papers that have already been published so far and encompass the use of ML algorithms in the HRI field. Thus, a total of 30research papers devoted to HRI were examined and analysed to give the most ML algorithms implemented in the field of HRI. Evidently, this study shows that the Neural and Reinforcement learning machine algorithms that are used mostly in the recent studies that have an interest in HRI use a machine learning algorithm with a supervised technique in a physical application. There are many challenges facing HRI using ML algorithms, which reduce the use of other ML algorithms such as deep and SVM learning algorithm. Unfortunately, these challenges limit use in social and mobile applications.

Keywords: *Artificial Neural Network (ANN); Deep Learning (DL); Deep Neural Network (DNN); Human-Robot Interaction (HRI); Long short-term Memory(LSTM); Machine Learning (ML); Recurrent Neural Network (RNN); Reinforcement Learning (RL); Support Vector Machine (SVM), Conventional Neural Network (CNN).*

Received 4 July 2024; Accepted 15 November 2024

1. Introduction

In this modern era, robotics is a multidisciplinary field of science that deals with intelligent machines, well known as robots, which are capable of performing target tasks autonomously or semi-autonomously [1][2]. Robots can be as basic as machines that follow a set of instructions, or as advanced as AI-powered systems that can learn, adapt, and work with people [1][2].More advanced robots can realize their surroundings, notice patterns, and make decisions on their own, without needing help from people. This is what makes them different from regular machines [1][2][3][4]. Beyond any doubt, the invention of this science rock the world and it is considered an IT boom that accelerates the globe and rapidly revolutionizes the landscape of services technology[1][2][3][4]. Beside it makes a significant qualitative changes in how individuals work and live, this science have benefitted individuals as well as the entire societies [1][3]. In the broader sense, the new phenomenon of this science ranging from infrastructure designing of robots to how human operator deal and interact with robots through the different well-designed applications and innovative services [1][2][3].

Back and forth, the ultimate aim of Artificial Intelligence (AI) is building these robots[2]. The philosophy behind designing of reliable robots' user interfaces is to fuse the foundation of well-stocked Knowledge inside the Technology [2].Consequently, there is a need to develop software that mimics human interaction in knowledge and technology, capable of generating either a single solution or a prioritized list of potential solutions in a predetermined order for solving target problems[2].

Being more specific, this comprehensive survey, Human-Robot Interaction (HRI), as its name implies, is devoted to the design, identification and evaluation of the robotic systems that illustrates the communication between a human and a robot [1][2][3][4]. This involves all the designing process directly related to the manufacturing of the same robots and all the needed well-designed interactive interfaces that are used by human to control these robots [1].In this regard, considerable efforts have been made over the past decade on trying to humanize both the design and functionality behavior of social robots to increase their acceptance among humans.

As shown in Figure 1, one of the most famous humanoid robots is the Advanced Step in Innovative Mobility (ASIMO), which was developed by Honda between 2000 and 2018 as an forward-thinking step in creating a walking robot with impressive capabilities [1][5]. This company stated that the communication process takes various forms, as depicted in Figure 2. These forms enable the robot's system to interact more intuitively with the world around it by collecting complex sensor inputs that guide its movements and interactions [1][5][6][7][8]. As a matter of fact, depending on the workplace, this communication follows two models: "Remote interaction" and "Proximate interaction" [2][6][7][8]. The former one, “Remote interaction” is defined when the human partner and the robot are not linked to each other and are in a separate locations, i.e. distant interaction where the human partner is not physically near the robot, but controls it remotely from a distance, routinely through a user interface [3][7][9]. In contrast, the latter one, “Proximate interaction”, is defined when the human and the robot are closely linked together and share the same physical space (i.e. close, physical interaction)[8]. Since this model typically encompasses direct communication and cooperation between humans and robots, it demands more immediate and responsive interfaces [2][6][8][9]. However, the appropriate choice between the two mentioned models is based on the specific application and the level of physical presence required in the HRI systems. Not only that, both interaction models are essential for developing multipurpose robotic systems that can adapt to a wide range of cooperative environments and meet diverse user needs.

Fig. 1. Honda ASIMO Robot from 2000 through 2018 [5].

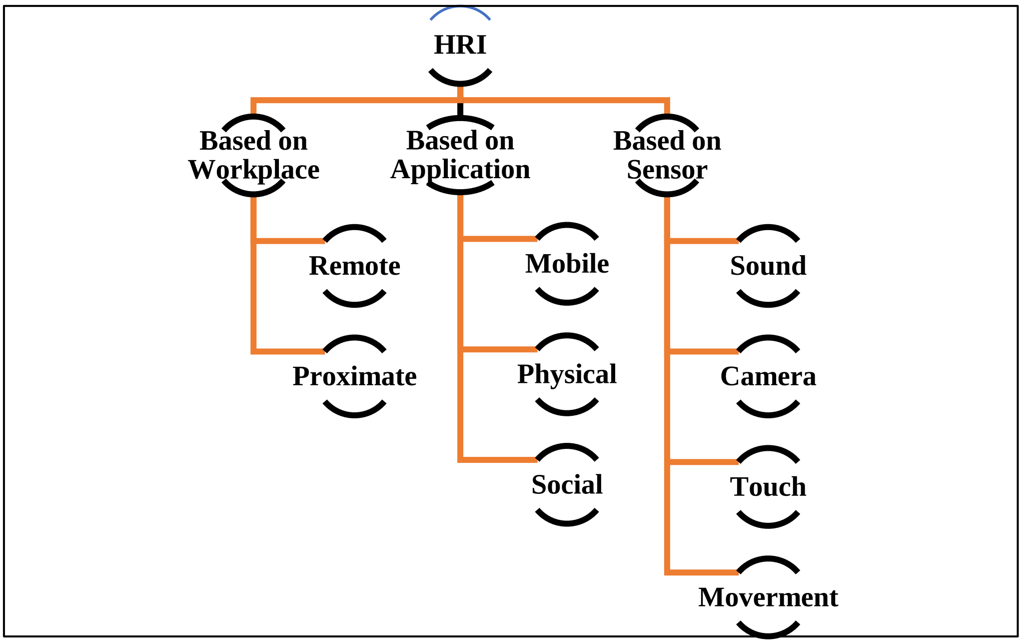

Fig. 2. Human-Robot Interaction (HRI)

Within these two types of interactions, it is appreciated to diagnose between the applications that require social, physical, and mobility interactions [10] which are already depicted in Figure 2. Furthermore, these types of interactions and the various HRI applications that determine the usage of robots are demonstrated in Table I.

TABLE I. IINTERACTIONTYPES AMONG HUMAN AND ROBOT

Application	**Remote Interaction**	**Proximate Interaction**
Mobile	Robot works as a remote operator.	The robot functions as a personal assistant.
Physical	Robot works as a remote processor.	Robots include physical interaction.
Social	Robot works as a comrade (i.e. acts as a helper, partner, or companion).	Robot includes emotive, and cognitive of interaction.

It is crucially important to mention that when the cloud-hosted services are used, the “Remote” model can be accomplished by using attractive high-agile services and well-built interactive interfaces which are really no more than web-based browsers accessed and shared remotely via the Internet [3]. On the whole, this cloud-like service is what is commonly known as Software-as-a-Service (SaaS); however, it is also sometimes referred to as “Cloudware” [3].

Indeed, this comprehensive survey studies the human impact when interact with the robot agent instead of a human [5]. Commonly, there are core distinctions among the different-used fields of HRI and robotics; since there is often some confusion about the difference between them, the following bullet points shortly clarify these distinctions. The robotics is related with the construction of physical robots and the various techniques in which these robots work in the physical domain [11].

The big notable difference is that the HRI is actually concerned with the ways in which intelligent machines (i.e. robots) interact with human within the social realm [12], while, on the other side, the robotics are primarily address with the creation, development, deployment of the physical robots and the ways in which these robots function in the actual physical world [11]. To be more exact, in mobile interaction, the robot is often considered as robot assistant. For instance, the robot assistant in the manufacturing will be accomplished through close interaction with human as to support their works, but not to replace them [1][13].

Accounting for the interfacing ways, the interaction may be achieved by several sensing channels within a collaborative work environments [1][5]. As illustrated in Figure 2, these channels include, but are not limited to, the following: touch channels (Motion Capture Systems) for sensing tactile input, visual channels (RGB-D cameras) for capturing images, and hearing channels (microphones) for capturing sounds[1][2][9]. Usually, the sounds and speech are received through sound sensors that might be converted into written text via pre-prepared dialogue structure [9][14]. On the other hand, some possible interactions are occurred by the robot which may use touch sensors to detect emotional gestures such as close proximity, punches, hugs, and caresses [14]. Furthermore, visual interaction may take place through a screen, usually a touch screen, allowing a substantial of information to be conveyed, whether visual or on the form of text [1][15]. It is important to realize that besides that sensors are equivalent to the human senses, the advanced ones can monitor some other sensations for which human couldn't sense [16][17][18].

The well-known dataset called LISI-HHI, which stands for Learning to Imitate Social Human-Human Interaction was introduced to assist researchers working on multimodal learning by providing a detailed examination of how people act together [19] This high-quality dataset is all about recording real-world interactions between two people in different communication circumstances. This dataset is a creative one because it uses many types of information, famous as "modalities". Advanced tools are used by the authors to gather the data like the following: data about where people are looking, how they move, what they resemble and even what they are saying; some if these tools are depicted in Table II. To be truthful, their effort make an important step forward to for teaching robots how to understand and imitate human behavior, which is a big challenge in the field of robotics. [19]

TABLE II. ADVANCED INSTRUMENTS USED FOR MULTIMODAL DATA COLLECTION [19]

Interactions Objective	**Tool**
To sense and record the sounds of their voices.	Microphones
To see where people are looking.	Eye Trackers
To capture both regular video (i.e. RGB video) and depth information.	RGB-D Cameras
To track the movements of the people.	Motion Capture Systems
To measure distance.	Ultrasound Sensors
To detect nearby objects without physical interaction in order to evade collisions.	Proximity Sensors
To capture social interaction dynamics and contextual behavior.	Multimodal Sensing
To analyze social interaction from a combination of visual, auditory, and motion data.	Integrated Sensor Systems
To capture facial expressions and body gestures.	Face and Gesture Detection Systems
To detect touch.	Force and Haptic Sensors
To focus its vision.	Adaptive Optics Sensors
To analyze speech and linguistic patterns in social interactions.	Speech Recognition Systems

As a consequence, this scholarly-research paper is grained down into seven sections. After this current section reviews some fundamental concepts and terminology that form the theoretical background, the reminder of this paper is organized as follows: Section 2 states the research problem statement and then tend defined the objectives of this paper. Section 3 explores ML algorithms that may be used in HRI. Section 4 is where the actual work begins; it presents the research methodology. Section 5 surveys the literature to look at the recent HRI using ML studies. While Section 6 discusses the most critical challenges facing the acceptance of robots' presences in our life and highlights the findings, providing a detailed analysis of how they align with the established methodology and criteria, Section 7 discusses the key challenges facing HRI in the future and proposes potential solutions. The conclusion of this research is presented in the final section, Section 8.

2. Research Problem and Objectives

Evidently, one of the key challenge in HRI is to design and model interactions between robots and humans in a way that is both flexible and adaptable to various dynamic environments and tasks [2][7][12]. To state the truth, rather than relying on traditional programming techniques, more dependable learning processes can be effectively employed to address challenges some of the challenges that are existing in HRI, such as ensuring that robots continuously improve their knowledge through repeated tasks and never stop learning[2][7][11]. In other words, robots can learn autonomously without calling for specific programming for each co-manipulation task in HRI[1][2][7].

To this end, machine learning (ML), a broad and essential field of AI, plays a critical role in HRI by enabling robots to learn and adapt to physical interactions with humans[2][11][16]. Through ML, robots can understand and react to human actions in real-time mode, improving their ability to cooperate and assist in dynamic environments [2][11][16][17]. Due to this, there was certainly a necessity for proposing a wide group of machine-learning methods and techniques in the past decade [3][16][17].

To delve deeper into the arguments that have elaborated so far, the central purpose of this research study is to introduce the latest ML methods and techniques that are used by robotics in the rapidly growing field of HRI, to identify and recognize the different shapes of human-robot interaction, and, ultimately, torise their approval among humans.

3. Machine Learning (ML)

To ensure that this research study is self-explanatory, this section establishes preliminary knowledge of the background pertaining to the concepts of HRI paradigm and draws a comprehensive image for both the current and future landscape of robotics research.

Since any consistent model or system should be fully trialed, the ML system needs a well-defined dataset, called training dataset, where the robot can learn how to acquire the data to create adequate creative knowledge [13][17][20]. It is essential to highlight that right training of the machine learning system is key to developing a reliable and trustworthy model[17][18].After the essential proper-sufficient training process or as a so-termed learning process, the model (i.e. system) should be examined on a great number of candidate samples[17][18]. The well-built trained dataset may encompass data that may be derived from the sensor circuits and has usually been manually defined and marked by humans[17][18]. An illustrative example on this when is the robot captures images of human faces from the sensors of a camera that gather this visual data and then transmit them to a controller to be further processed [16][18]. Ultimately, the controller interprets the emotion of that human and categorizes it into three classes: "neutral," "smile", or "angry" [1][2][20].

Going further, the robotic sensor may be used to detect as signals the surrounding environmental conditions such as sounds, images, video, temperature, humidity, pressure, movement, light level, radiation, and more [18]. These signals are forwarded to the robot controller to detect, interpret, and analyze the applicable behavior and, in many cases, responding accordingly and taking the relevant behaviors and the justified actions [18]. Figure 3 depicts a robot, or as so called mechanical man or humanoid, that is outfitted with a variety of sensors and communication elements, including visual, audio, ultrasound, proximity, and several others. These components have a great effect in helping this humanoid robot to mimic human by sensing and interacting efficiently with the world around it. To further simulate human-like communication and environmental understanding, a humanlike may be equipped with sensors for smell, taste, pressure and humidity measurements, and electrical sensing. It is important to note that robots typically require a substantial amount of information to operate effectively [3][18]. This extensive data is crucial for ensuring that they function correctly and can perform their objective tasks as intended. Without this comprehensive input, the robots may struggle to execute their functions accurately or efficiently [3][18].

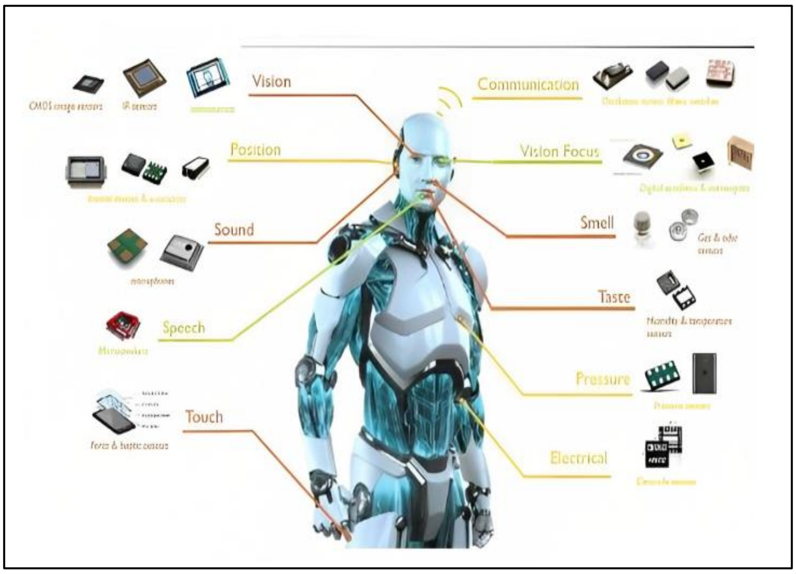

Fig. 3. Humanoid Robot's Sensory and Communication Components[21]

Computer vision, on the other hand, is one of the vital fields within HRI[1][22][23]. It interprets a series of 2D images captured over time, particularly when dealing with video data. Furthermore, Computer vision allows helping in self-localization, detection, mapping and tracing of individuals [22][23][24].

As discussed, beforehand, the relevant data gathered by the robotic sensor is processed and translated to a more appropriate interpretation in order to discover the core features [20][25]. To be exact, this is what it's so-called extraction feature in machine learning (ML) [4][16][26]. In essence, there is undoubtedly a wide group of algorithms designed to extract features directly from the raw input data collected by sensors [27]. These features are arranged in a vector, typically called a feature vector or embedding, which is a row of numbers organized for further processing[4][16][26]. Not surprisingly, some experts and researchers commonly evaluate their data sets manually to define the most significant features. Moreover, the training process of an HRI model is typically preceded by a pre-processing phase for signals, called the preliminary processing phase[20]. This phase involves removing noise (i.e. filtering noise) from the signals before assembling the feature vector[20].

Many recent machine learning (ML) algorithms are used by robotics in HRI based on Deep Learning (DL), Reinforcement Learning (RL), Support Vector Machine (SVM) and Artificial Neural Network (ANN) [17][18]. It is essential to mention that each one of these stated algorithms has its remarkable strengths as well as potential drawbacks that are important to be considered[11][13][16][28][29]. The hierarchical relationships between the different types of ML algorithms are clearly depicted in Figure 4; these algorithms can be categorized into various types into. The figure highlights that ML algorithms are grouped into four subgroups: Supervised Learning (SL), Unsupervised Learning (USL),Semi-Supervised Learning (SSL), and Reinforcement Learning(RL).SL is used when training data sets with known outputs and inputs (i.e. only labeled data) is used, and is therefore very beneficial in industrial processes where quality control is sought at work stations [18][20][30]. On the hand, the USL is especially proper for the cases in which training with dataset are unknown inputs and outputs (i.e. training with only unlabeled data);this allows the model to organize information or find out patterns within the data set by itself without being directly taught by human partner [18][30][31][32].In the cases where acquiring large amounts of unlabeled data is easier, more timesaving, and/or cost-effective, a collective fusion of small amount of labeled data added to a large amount of unlabeled data is formally referred to as SSL. It is crucially important to mention that the learning process in this subgroup (i.e. SSL) is obviously guided by this relatively small amount of labeled data.

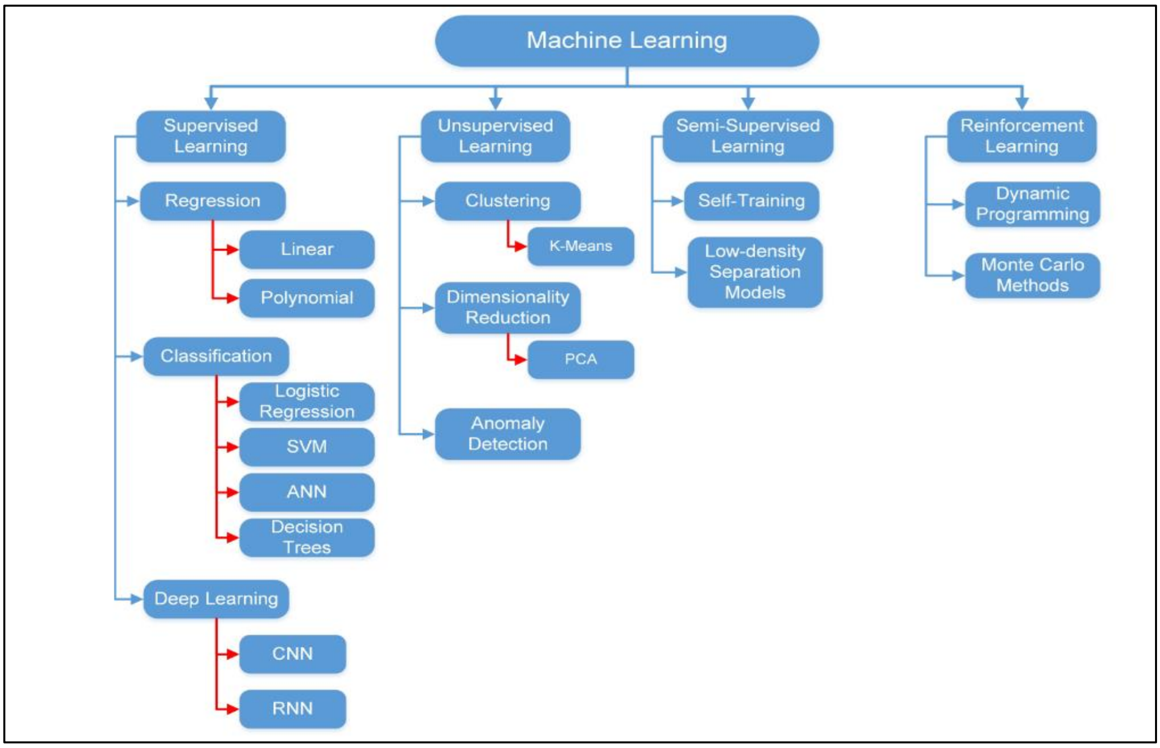

Fig.4. Machine Learning Algorithms[33]

Broadly speaking, classification, which is a type of supervised learning, is a major machine learning technique frequently applied in HRI to enable robots to recognize, interpret, and respond to diverse human behaviors and inputs [28]. Classification is used to predict the most accurate label, also known as class, of a given input data. Classification assigns a class to a data point based on the training data [34]. Returning back to the example of an image captured for a human face by a given camera, the classifier decides the most accurate emotion of that human as one of the three classes: "neutral," "smile" or "angry" [18][35][22]. Back-and-forth, HRI wants ML algorithms to manipulate precisely any piece of information that has never been shown. Consequently, machine learning can give rise to algorithms used for face detection technology for tracking human by robot [23][24][1]. The robot is equipped with a vision system that is able to detect human faces through its camera input and uses these data to identify, locate, and, accordingly, track the movements of a target person in real time mode

[23][24]. The tracking activity is continuously adjusted to align with the target's position as it moves[1][23][24]. Whenever this occurs, the algorithm improves the data that it will be trained enough, but it may works badly when it comes to new problems [23][18][20].

For the sake of reduce the human intervention during the training process, Deep learning (DL) is invented which is a ML for training large artificial neural networks (ANNs) [4][16][26]. The main idea of using DL is that it does not require manually careful extraction of the features; this is managed by using multi-layered neural networks, that have huge number of connection weights (parameters), to automatically learn representations of data, hence the word "deep" is used [17][18]. This stacking multi-layered networks allows the model able to remember things for a long time. Additionally, deep learning (DL) is highly effective in enabling robots to adapt and progress on-demand over time by continuously learning from new input data, further enhancing their ability to perform complex and large-scale tasks with minimal human intervention [16][17].First and foremost, since the DL algorithms are non-linear, it is considered as an ideal choice for the robotic applications; this is because it can learn features directly from the data [16][36].Despite the fact that Deep Neural Network (DNN) algorithms do not require human experts to extract patterns from datasets and possess a high level of self-adaptability, they have a significant drawback: they require large volume of diverse data to achieve extraordinary performance without needing of human manual intervention[16][17][26]. This practical limitation draws the reason behind that DNN algorithms often require significant processing power and memory that is greatly associated with network architectural complexity, leading to high computational costs[16][17][26].This limitation is proportionally scaling with the problem size, in both memory and time. Furthermore, there may be issues related to data availability, which can make experiments less practical in applications where data are sparse, so hard to come by, or very costly to obtain[16][17][26].

Another advanced and interdisciplinary area of ML paradigms is reinforcement learning (RL). The robot in RL gets rewards, indicating positive or negative progress, from its surrounding environment to see how well it does the target task effectively [37][38][39][40][41]. The robot tries to do better performance by getting more positive rewards for good actions [42][43][44]. This is attained by obtaining more experience and directly engaging with the robot's surroundings. It should also be noted that, in order to help the robot to learn better, humans can provide it with the knowledge in the form of suggestions. However, the RL potential drawback is the robot may not be able to make accurate relations between the actions selected and the states that are observed that lead to greater rewards[40][42][43][44]. What is more, a lot of memory (i.e. memory-intensive) is requested to store the values for each state[37][38]. To tackle these potential problems, we can avoid saving every combination of states and actions[45][38][41]. Instead, these combinations of pairs, states and actions, can be calculated on the spot when necessary [45][41].

To conclude this section, ML can be employed to develop robot behaviors, enhance robot perception, vacillate multi-robot interaction, and assist the robot in learning from interaction with the human[37]. HRI utilizing ML has recently attracted significant attention among scientists and researchers across different research foundations and universities around the world as a tactic to capture and encode valuable robot behaviors, vacillate robot training, and enhance perception [10].

4. Research Methodology

The main objective of this scholarly research paper is to provide a state-of-the-art review that covers a wide range of studies, from classical to novel approaches, on HRI using machine learning (ML). This research uses a set of previously published research papers selected from various journals including IEEE Explore, ACM, Scopus, and others. To this end, these papers highlight ongoing efforts to tackle safety concerns in HRI using ML, with a particular focus on collision detection and avoidance in cooperative human-robot settings. These research studies are, in turn, for enhancing safety measures, designing adaptive safety mechanisms, and emerging collision avoidance strategies. Consequently, this research paper examines how ML techniques can be combined to enhance safety measures in HRI. From a criteria point of view, the set of research studies is selected by researching on Google Scholar using some keywords related to HRI such as: "Interaction Human and Robotics", "Collaborative Human and Robotics", "Human-robot Interaction", or "Interaction Human and Robotics using Machine Learning". Besides that, the research comparison factors that are used in this added-value research paper are the type of robot sensor, HRI application, and the type of ML algorithm that is implemented in HRI.

Related to the previously mentioned research methodology and criteria, the following section presents some research studies that are relevant and have contributed to the field on this topic.

5. Survey of Research Papers on HRI with ML

The main focus of this core section of our survey is to look at and review recent studies on HRI that utilize ML. A total of 30 different research papers, all involving ML algorithms in the context of HRI, were reviewed, examined, and analyzed. As a consequence of that, the analysis, along with comparison, and discussion of these research studies will be presented in the next section.

To develop a new and efficient way to evaluate how humans and robots work and cooperate closely together to accomplish common tasks while minimizing collision **issues**, the research work in [46]is introduced. This study focuses on recognizing potential collisions in a shared workspace and, in turn, improving the safety measures in the course of HRI where the robot works alongside the humans as an efficient well-coordinated teamwork that understands and reacts appropriately to human actions. For achieving this goal, the authors of this paper employee three different types of ANNs:

- Feedforward Neural Networks (FNNs): This type is the traditional one of ANNs, generally used for pattern recognition and classification tasks. Since there is no any feedback loops, the data move in only one direction, from the input to the output.
- Recurrent Neural Networks (RNNs): Since their architecture is based on loops that permit information to persist for long time, these types of networks are used for processing sequential data. RNNs are often applied in tasks involving natural language processing or time series data.
- Convolution Neural Networks (CNNs): Since this type is particularly useful at automatically and adaptively detecting patterns in visual data, it is usually utilized in image-related applications such as video analysis.

These three ANNs helped in recognizing different objects within the workspace, filtering out unnecessary data, and tracking the movements and interactions of both the robot and the human operators. By doing this, the system could better predict and prevent possible collisions between them; therefore ensuring smoother and safer collaboration.

It is vital to ensure safety to avoid accidents and coordination problems and to better determine the relationships between data points, the research study referenced in [47]introduced a Deep Metric Learning (DML) algorithm, which is a specialized subset of metric learning, and metric learning is itself a subfield of the broader field of machine learning (ML).Evidently, the philosophy behind proposing this smart algorithm is born out of the necessity for developing a distance function or similarity measure that can effectively compare different data points and explore different relationships among them. It extends the old-fashioned concept of metric learning, which focuses on distance functions, to a broader framework that studies non-linear, complex relationships between data points. From a different point of view, DML is a promising methodology for improving human-robot collaboration, particularly in industrial communities; it fuses the goal of metric learning with the power of multi-layered neural networks, making it highly effective for specific tasks that require an automatic understanding of complex relationships between data points. Furthermore, this DML-based method enable robots to anticipate human actions and take preventive measures to enable robots to better interpret and predict human spatial and movement. The DML algorithm was verified, and evaluated across different settings. As a result, this algorithm improves HRI safety measures to mitigate or prevent potential hazards that may occur when humans and robots work closely together.

Another research study on DML is provided by S. Li [48], focusing on its key principles and deep learning architectures. In this paper, DML is used for examining some so critical components, such as optimization strategies, loss functions, architecture designs, discovering hidden relationships and patterns within data, to mention but a few. These components cover a wide range of applications, including image retrieval for content-based search, few-shot learning for tasks with incomplete labeled data, handling imbalanced data as a learning algorithm for embedding spaces based on the relationship between data points. These components cover a wide range of applications, such as finding images based on their content (i.e. content-based search), learning from a small number of examples when not all data is labeled, and managing imbalanced data as a learning algorithm for embedding spaces according to the association between data points. Furthermore, the last talked algorithm is measured by the similarity features where similar data points are put close together and dissimilar are put farther apart. An example of that is recognizing faces for identifying individuals where different images of the same person's face are placed close together and, on the other hand, images of different faces of different persons are placed far away from each other. For this well-defined reason, a learned distance function is used to guide this categorization or arrangement in order to figure out what looks the same and what doesn’t. So, different images of the same person’s face are placed close together, while images of different people’s faces are placed farther apart.

It's important to note that DML should not be confused with matrix learning, which uses matrix factorization techniques to predict missing data or recover corrupted values based on observed information. This is often also called matrix completion or matrix factorization. It is fabricated to know a representation space where the distances between any two data points may reflect some significant relationships between them, such as similarity or dissimilarity. By learning or assuming distance metrics from existing data this approach, in one way or another, is frequently used to roughly estimate the missing entries in a given matrix. In other words, this approach is used to fabricate new data based upon the already existing data.

Based on the idea that people's knowledge arises over time when they recall what they have previously learned in past lessons and, conversely, declines when they are unable to remember (i.e. forgetting) past information, Ayub et. al. [49]discuss the concept of accumulated knowledge, or as so-called continual learning, through repeated interactions. Using this interaction rule, the authors demonstrate how a robot can develop a warm and natural relationship with humans over time. The robot used in their experiments is "Fetch mobile manipulator" and they prove that their system can recognize objects and, on top of that, it doesn’t stop learning -it continuously improving its learning skills over time and, in turn, the gained knowledge is increase dafter each repeated session of interaction. To help robots keeps improving its learning skills from humans over time, their system uses three different continual learning models:

- Cumulative Learning Model (CLM): The robot in this model continuously learns new things over time while storing everything it previously learned. It keeps all the newly acquired information along with its prior skills, allowing it to improve and adapt through repeated interactions.
- Adaptive Learning Model (ALM): Based on feedback from people and situations, the robot in this model works in a different way than the previous one by adapting its behavior as it repeatedly engages with humans. This model resembles how people change their habits by revisiting them periodically, based on past experiences.
- Selective Learning Model (SLM): To avoid being overwhelmed in the long duration by storing too much stacked data over each other, the robot in this model doesn’t preserve all the prior information it receives. Instead of that, the system dynamically selects the most important or useful information to keep and remember, while discarding unnecessary or fair-bit less relevant information.

In the field of Intelligent Transport Systems (ITSs), Munguia-Galeano et. al. [38]introduced a novel and valuable method to increase robot cooperation with humans by using a method called Contextual Q-Learning (CQL), which is based on Reinforcement Learning (RL). It helps the robot to reduce the number of actions it needs to choose from (i.e., reducing the possible alternative soft he action space size), thereby increasing its learning speed and making it faster at figuring out what to do. A practical application of this method is to help the robot use information gathered from its surroundings to consciously adjust its actions based on the updated situation. This improves the robot's ability to perform specific tasks like picking up, moving, and releasing objects when working together with people during human-robot interactions (HRIs).From a purely practical standpoint, this study showed that this method helps the robot to learn quickly to resolve many real-life problems with a high success rate.

The paper referenced in [50] discusses how, with the help of Support Vector Machine (SVM),a robot can interpret and react to human operators' actions in real time. In the context of real-time vision-based Human-Robot Interaction (HRI), a robot can identify facial expressions, detect a person's gender, and recognize real-time facial gestures. In all cases, the dataset includes facial landmarks categorized as “sad,” “angry,” “smile,” “surprise,” and “normal,” which are used to train the SVM to generate a classifier.

In addition, the paper referenced in [51] explores the use of neural networks for classifying human-robot contact situations. It focuses on how a neural network can be effectively trained on measurement data gathered from various contact scenarios between a human subject and a collaborative robot (cobot), such as the ABB YuMi robot. For collaborative robots, the paper emphasizes that classifying contact situations is vital for developing safe HRI protocols. It is vitally important to note that physical contact situations are labeled as “collisions”, while anticipated interactions are labeled as “interactions”. Moreover, the system is designed to conclude whether contact occurred on the robot's upper or lower arm.

Another study aimed at solving the problems faced in ITSsby using machine learning (ML) was conducted by Wang et. al. [34]who suggested a system called ML-ITMS to make traffic monitoring safer and more reliable. ML-ITMS stands for Machine Learning-Assisted Intelligent Traffic Monitoring System. The main objective of this system is to predict and manage traffic flow more accurately and efficiently. HRI helps address important concerns for both consumers and service suppliers in the transport system. The types of sensors used in their study are summarized in Table III, which provides a clear summary of each sensor's purpose and how it contributes to the overall system.

TABLE III. SENSORS USED IN ML-ITMS [34]

Interactions Objective	Tool (types of sensors)
To captures visual information.	Camera
To track the movements of the objects by measuring deflection, displacement, movement, and vibration that are not visible to the human eye.	RDI (Remote Device Interface)
To detect and measures distance and speed.	Radar
To measure distance to objects.	LiDAR (Light Detection and Ranging): It is a remote sensor that uses laser light.
To monitor environmental conditions (e.g., temperature, humidity, wind speed, and pressure).	Environmental Sensors
To capture sound and vocal commands.	Microphone
To provide tactile feedback and measure forces.	Force Sensors
To provide movement measurement and direction.	Motion Sensor
To measure and record the torque on rotating systems.	Torque Sensor
To measure location and displacement.	Position Sensor
To measure speed of robot movement.	Velocity Sensor

Ottakath et. al. [24] proposed an innovative model using deep learning for social distance measurement and mask, which can be integrated into both stationary robots and mobile ground devices. Their proposed approach aims to develop safety and ensure adherence to health guidelines across diverse robotic platforms.

Malik et al. [52] examined an industrial case study to address the complexities of collaborative production systems by exploring the potential of utilizing digital twins. Their study investigates various forms of digital twins throughout the lifecycle of a collaborative robot system. By analyzing these digital twins, the authors demonstrate how they can enhance system performance, improve integration, and streamline operations in collaborative environments.

Yan et al. [53] designed an improved, smaller, and lighter version an optimization techniques for detection and harvesting apples. This model is called "YOLOv5s", where the abbreviations "YOLO", "s", and "5" stands for "You Only Look Once", "smaller", and "the 5th version", respectively. YOLO is a well-known AI algorithm that uses Conventional Neural Networks (CNNs) for object detection. Furthermore, extensive real-time experiments in detecting apples accurately and quickly are conducted by the authors to validate the system's effectiveness. Both high speed and precision accurately on the way of recognizing and harvesting apples are noted, and, consequently, this work has significant implications for advancing automation in agriculture. This apples' lightweight-solution can be extended for cultivation and harvesting other plants. Not only that, this deep learning-based object detection algorithm for this agricultural solution can be extended for developing more well-organized and intelligent picking robots for real-time detection of other objects. From an alternative perspective, this applicable and practical solution is an important step forward to in the advancing of other practical applications where real-time object recognition is critical. This is particularly relevant to the security systems, autonomous driving, military manufacturing, and other robotics-related domains. A further point to note is that, in addition to YOLOv5s, there are three more advanced models with slightly higher accuracy: YOLOv5m, YOLOv5l, and YOLOv5x, where the notions 'm,' 'l,' and 'x' stand for medium, large, and extra-large, respectively.

Barstuğan and Osmanpaşaoğlu [54] made a robotic hand with a 3D printer. It was attached to a robot arm, and each finger had its own motor to move it. People wore special gloves with sensors that told the motors where their fingers were. The glove sent messages to the robot very fast, so the robot could move more smoothly. This helped the robot hand pick up (i.e. hold or grasp) move and release the objects. The robot arm was moved using pictures processed by a computer. The authors used two versions of the YOLO program, YOLOv4 and YOLOv5, to see which one was better at finding the gloves. They also used 5G communication to make the robot faster and more responsive. This project showed how robots can be controlled by humans in a simple way. In other words, this project tried to make robots better at understanding people’s movements in real time. As a consequence, their project can be useful for robots in places like factories or hospitals.

Instead of using a big and expensive robot, Ahmad et. al. [55] proposed a small robot system for a medical condition called Twin-to-twin Transfusion Syndrome (TTTS). Their research study looks at how to automatically find the placenta's position using pictures from a very small camera used during fetal surgery. By looking at just one picture taken during fetal endoscopy, the authors of this paper successfully used deep learning, especially CNN, to guess and manage where the placenta is. Since their approach uses a simple, single-camera setup, this makes it more accessible and reasonably less costly than complex systems. Generally speaking, their CNN model was well-educated enough to accurately estimate how the placenta is located and oriented from just using one camera image during the surgery. Undeniably, this research study is important because it helps in buildinga starting point for making fetal endoscopy easier and partly automated in the near future. This means that in the years to come, physicians or surgeons might be able to use tools and techniques from this study to perform these so complicated procedures with less manual effort and more knowledge. In a wider sense, this could lead to safer and healthier experiences for both obstetricians and patients during fetal surgeries.

A new method of teaching robots, called Interactive Reinforcement Learning (IRL), is proposed by Akkaladevi and colleagues [39]. Their method helps robots learn how to work together with humans in a complete assembly process. The learning process of their method contains two major steps. In the first step, called Task Modeling, the robot is well-trained enough to understand the basic requirements of the learning tasks that need to be accomplished. These tasks are part of a well-organized list known as task-based procedures, which include various guidelines and options for these tasks. The human, referred to as the operator, plays a role in feeding the input data into the Reinforcement Learning (RL) network and helps identify the precise actions the robot should take accordingly. In the second step, called Reinforcement Learning Implementation (RLI), the robot's actions are optimized to the greatest possible extent based on the feedback of the cooperative environment. The human partner provides positive or negative feedback, known as a reward, depending on the degree of collaboration and the time required in completing the task to be accomplished. This step is repeated iteratively to allow the robot to improve its performance. It continues until the human operator is entirely convinced that the robot has become effective and suited to work in a real-time environment, or in so-called real-time HRI scenarios.

For the purpose of enhancing teamwork coordination between humans and robots and build long-lasting trust in HRI, Chen et al. [56]created a special kind of computational trust-aware model that focuses on trust that arises over time. They designed a method to help the robot learn from its experiences. This method is called Partially Observable Markov Decision Process (POMDP). Over time the robot becomes better at making decisions and, in turn, adjusts its actions according to the trust feedback that are defined by the human partner. This means that when the human gives more trust to the robot, the robot can work better and become more effective in working alongside humans.

When both the human's and robot's arms have similar stiffness levels, they can work together more smoothly and efficiently as a tuned teamwork. Based on this ground, Chen et al. [57] outlined a well-built method to discover what the human intends to do with his/her arms when working beside a robot. The authors use a special device, called Myo armband, for measuring the stiffness of the human arm and compare that with the stiffness of the robot's arm. The word "Myo" is derived from another related Greek word "myos", meaning "muscle," and it is used now in relation to various muscle activities associated with modern technologies. Since this armband is worn on the human's wrist, it recognizes the different positions and movements of the wrist and decodes it into well-formalized control signals. Consequently, what the human intends to do can be detected by this device. Then the authors apply one of the ML neural learning algorithms which is used to help the system better recognize the wrist configurations. To conclude this talked study, their method has a significant impact in advancing the wheel of the collaboration between humans and robots to become more effective and, in turn, in making the robot work alongside the humans as part of an efficient, well-tuned team.

A new way to control how a robot reacts to forces during HRI is presented by the research study of X. Chen et al. referenced in [58].Their new method is based on approximately calculating the stiffness of the person's arm and then adjusting the robot's control parameters to correspond to this estimated stiffness. After estimating the human arm's impedance, a Linear Quadratic Regulator (LQR) is used to calculate the robot arm's admittance model, so it matches the human arm's impedance. With this adjustable control, the robot arm can work smoothly with the human arm by responding to the force applied by the human hand. In other words, as the robot keeps adjusting, it can work more smoothly and naturally with the person, making teamwork easier and more effective. This helps them work closely together more efficiently. A neural network is also used to handle unknown movements and ensure the controller works well. At the end, tests were done to check if this method is enough effective.

"Sawyer" is a type of collaborative robot, better known as cobot, created by Rethink Robotics, mainly to work alongside with people in industrial settings. It’s often used in factories, but because it’s flexible and can do complex tasks with human partners, it's also popular in research on human-robot teamwork. The authors in [59] presented a cognitive system that helps the robot coordinate actions and decisions during joint tasks with humans. This system uses connected neural networks, which function like small local systems with specific tasks. The authors tested this cobot in a real construction project, where it worked alongside a person. At each step, the robot decided and verbally announced which part to assemble next, then took the right action to put it in place. The two-dimensional Action Execution Layer helped show both the objects and the actions together. The outcomes showed that this cobot made good decisions, even in different work environments or when some parts were missing.

The authors in [40] described a system that uses reinforcement learning(RL) to help humans and robots work closely together as a teamwork. This system is made to work fast and lower the chance of errors, helping to complete predefined tasks more quickly. It learns by itself, figuring out both how to look at things and how to make decisions at the same time. The system was tested on a packaging task and showed that it could help humans and robots work closely together better than other methods that use guided learning, where seeing and deciding are learned separately. Two key advantages of this approach are that it avoids the need for detailed labeling of movement data and allows learning to happen in real time.

The research paper outlined in [60] suggested an emotion recognition system for a humanoid robot. This system (i.e. robot) is outfitted with a camera to capture facial images of users, enabling it to identify their emotions and respond accordingly. The emotion recognition system, utilizing a Deep Neural Network (DNN), is well-trained enough to recognize six fundamental emotions: happiness, sadness, fear, disgust, anger, and surprise. Furthermore, this proposed system works in a sequence of four phases that form its roadmap framework. First, a Convolution Neural Network (CNN) is employed to extract visual features throughout the training process, or as a so-titled learning process, conducted on a vast dataset of static images. Next, a Long Short-Term Memory (LSTM) recurrent neural network is used to evaluate the relationship between changes in facial expressions across image sequences and the six aforementioned fundamental emotions. Third, the proposed model integrates CNN and LSTM to power the strengths of both techniques. Finally, the emotion recognition system's performance is enhanced through transfer learning, which involves applying knowledge gained from solving related but different problems.

To improve safety and efficiency, to enable the robot to automatically adjust the level of physical assistance, and to dynamically enhance and fine-tune real-time human-robot collaboration for smoother and more natural interaction, the research study referenced in [42]proposed a reinforcement learning (RL) controller that combines auxiliary control with variable admittance control. This combination lets the robot dynamically regulate its behavior according to the user's desired actions and basic requirements. This predictive controller aims to optimize the response time of memory networks (i.e., reducing latency) and anticipate human intention, particularly in environments where real-time flexibility is crucial, such as in sensitive factory settings that require close cooperation between robots and humans. This means that robots can respond more effectively to human movements, making collaboration safer and more efficient. Ultimately, the proposed adaptive impedance controller was validated using real-time joint torque and force sensors, indicating that it achieves smooth, low-effort operation with a minimum-jerk trajectory.

The paper referenced in [43]discusses an approach using a Model-Based Reinforcement Learning (MBRL) controller to support humans in performing teamwork tasks with robots. In this proposed approach, a set of Artificial Neural Networks (ANNs) is well-trained enough to learn and understand the way humans and robots interact closely together. These ANNs are, furthermore, designed to account for any uncertainties that may arise during this interaction and teamwork collaboration. The system can adjust stiffness and damping in real-time by using these learned models; this helps in reducing the physical effort required from the human. What's more, the well-stocked knowledge gained after these neural networks learn how humans and robots work closely together is then fed into a Model Predictive Controller (MPC), which helps make decisions on how to act. After this important step, a technique, called Cross-Entropy Method (CEM) is then used by the MPC. This technique, in somehow or another, helps in optimizing actions by selecting whichever the best one out of the possibilities. An important point to mention is that their system is continuously adjustable, making the collaboration between the human and robot easier and smoother for the human.

The research work in [44], reported that an adaptive impedance control system was created where humans and robots can work closely together and collaborate more effectively in accomplishing some given such as passing tools and jointly moving or lifting heavy objects. This system allows the robot to modify its motion and control its impedance parameters in real-time mode without needing any earlier knowledge about the task to be accomplished. Instead of relying on a detailed understanding of how the system acts and behaves, Reinforcement Learning (RL) is used to help the robot to compute and pick up the right set of parameters. These parameters are generally chosen to reduce costs related to the task’s goals. After the robot has learned the right parameters, it fine-tunes them further by considering any disagreement from the human partner. The system is designed to estimate the human's motion reference by using a simplified model of the contact dynamics and identifying the system's behavior through a process called system identification. This method helps the robot to actively contribute to the task while remaining flexible to changes in the environment or task itself. The authors of this research also presented impressive experimental results that demonstrate how well the robot performs using this method.

The authors of the research paper referenced in [61]introduce bilateral control as a method that allows both the human and the robot to jointly effect each other’s movements. They also introduce imitation learning as a machine learning (ML) technique in which a robot learns to perform certain tasks by observing and imitating human actions. By this technique, the robot watches pre-recorded demonstrations, registers the actions, and then attempts to redo the behavior autonomously or collaboratively with the human. Based on practical real-time feedback, they indicate how imitation learning and bilateral can be used together to design framework that is used to expand and coordinate the collaboration between both humans and robots in numerous learning tasks. Because their research study demonstrated that robots could adjust their actions in real-time mode to match human movements based on the prediction of human actions, their framework helps in making human-robot interaction smoother and more efficient. Furthermore, this real-time adjustment in collaborative settings enhances coordination that produce better teamwork, improving how well the human can work alongside a robot, minimizing interruptions or errors, reduces the cognitive and physical load on human operators, and improving the overall coordination between human and machine.

Since the Recurrent Neural Networks (RNNs) issued successfully in detecting patterns of the actions that happen over time, like human actions and movements, the study in [62] introduced a method using this type of networks to predict how humans move during some collaborative shared tasks. The authors' aiming this research paper is to bridge the gap between understanding human movements and making sure those robots can respond efficiently as human. By observing their actual practices and then predicting their next movements, the authors use Deep Learning (DL) to analyze how people move in assembly tasks and accordingly help robots to plan and take correct real-time actions. They empirically demonstrated that their proposed approach works well in an industrial environment by using engine assembly setting to help in forecasting what the human operator will do next and then tell the robot to respond accordingly. By employing this tactic, the robots can better anticipate and assist in performing the required tasks, enabling smoother collaboration between both robots and humans.

A new approach, called the Dynamic Neural Fields approach, is presented by the research paper introduced by Wojtak et al. [63]. Because the authors try to use calculations similar to that of human brains, their proposed approach helps the robots think and act like humans. Their approach contains a special key component called a neural integrator that works by collecting information from the surrounding environment. The authors tried to make this component work as the human brain and to function in a way similar to how the human brain works and acts. This component analyzes different possible actions that the robot could take. A dominant example of this is when a robot senses something by using its sensor where the robot's neural integrator studied the various actions that may be taken. The final selected robot's action is chosen after making a race between these actions looking for the best one. This racing of competing against each other to find the best one is called competitive dynamics and it similar to how the human reacts after sensing something and then decides to do a selected action, like, for instance, moving an arm or turning the head. Like that, the robot's neural integrator is connected with different groups of "neurons" to carry out this competition for picking up the most important or the strongest action between the available ones.

The research described in [32] introduces a framework that helps robots in deeper understanding what their human partner thinking or planning to do when they are both working closely together as a cooperative partnership. This framework uses two approaches to guesstimate and interpret human intentions, Unsupervised Learning (USL) approach alongside a probabilistic approach. The former approach helps the robot to learn patterns on its own by itself without needing direct instructions from human partners. The latter approach, on the other hand, is generally used by this framework to help the robot in estimating or calculating what the person's goals are or understand human actions. So the robot needs to make smart guesses based on the information it can see because it assumes that it doesn't always know what the human wants to do. Furthermore, this probabilistic approach can be used with social cues including eye contact, pointing, body language, and gestures. These cues help robots understand what a person means or wants. Overall, both of these approaches help the robot to better sophisticated understanding and cooperate with the human partner's intentions during cooperative activities that have some uncertainty, such as working together on a project or playing on the same game. In short, this proposed framework is increasingly important to build a better partnership between humans and machines.

The research study referenced in [64]introduced a way to improve actions in tasks where people and robots work closely together with constant physical contact. Their study emphasizes the importance of the work environment in these tasks. The method aims to improve ergonomics and efficiency. By considering the work environment, robots can better assist humans. This leads to safer interactions and higher trust levels between robot and human partners, as well as more effective collaboration between both humans and robots.

The authors of [65] demonstrated a neural computational framework based on gradient optimization of the robot's target state. Their framework includes Convolution Variation Autoencoders (ConvVAEs) and RNN with LSTM architecture. Their setup learns to connect target images with actions the robot should take. It helps robots change their goals automatically. This makes human-robot teamwork more effective by allowing robots to understand and respond to their surroundings better.

As described in [66], the research introduced a method to watch how tired people's muscles get when they work with robots. A ML program learns how the strength of a person's muscles relates to how much power is used in their arm. The study looks at how people and robots can work better together, and, hence, HRI is enhanced. A detailed model of muscles and bones, referred to as musculoskeletal, is used to reach the goals of this research study. This model helps in understanding muscle tiredness during tasks to be accomplished. In the end, the approach seeks to improve the comfort and efficiency of users while interacting with robots.

In the research study mentioned in the reference [67], the authors came up with a new fuzzy-controller method that combines two important ideas in order to help human workers perform heavy industrial tasks with more safely and efficiently:

- Fuzzy-impedance control: A fuzzy controller is a type of controller that works like how people think and make decisions, handling uncertainty and using approximate educated guesses instead of exact answers. It's often used in robots, factory machines, and home devices, especially when creating a perfect math formula is too difficult to solve or can't be done. On the other hand, Fuzzy-impedance control is a technique that helps machines or robots move smoothly and safely, even when they are handling parts or objects with uncertain or changing weights.
- Safety rules: These built into this method are designed to keep operators safe while they work in companion with robots.

By integrating these two ideas, the approach not only enhances operational efficiency but also prioritizes worker safety. What's more, the combinations of fuzzy-impedance control and safety rules helps ensure that everything runs smoothly, reduces the risk of accidents, and improve the interaction between humans and machines. Their new method is particularly useful in big industrial settings where people often have to handle large and heavy parts that they might not be familiar with. By using this method, operators can get better assistance and control over these parts, making their work easier and safer.

Sometimes there may be some differences between how robots and people think about some tasks that need to be accomplished, even when they are both presented with the same information or circumstances. For this well-defined reason, the research study referenced in [41]proposed a way to compare how an autonomous system, like a robot, and a human partner understand tasks differently and figure out why they may not see things the same way. Then, this said study also analyzes, investigates, and compares what might happen because of these variations in perception and, for that reasoning, allows users to give their constructive guidance and feedback to guide the robot in making model adjustments and enhancements. This study combines two leading concepts: Reinforcement Learning (RL) and explanation-based coaching. The latter concept is mainly used by the system to give clear and simple justifications for its suggestions, making it easier for human partner to understand. On that account, figuring out this mismatch between the two, humans and robots, can significantly help in understanding why it happens and what might go wrong, hence improving their decision-making.

Finally, to wrap up this section, the study mentioned in [68]introduced a special type of neural network called a deep Long Short-Term Memory (LSTM) network. By analyzing patterns found in the stored data, this type of network is designed to expect what a person plans to do in long-run without being directly told. The authors used several LSTM layers stacked on top of each other, where each added layer contributes in making the model retaining more information over long periods of time; otherwise, this information may be forgotten too quickly. This arranging in a way or another has a great effect in allowing their model to solve the difficulty in memorizing more information and, hence, their model becomes better at understanding the complex patterns and relationships found in the data which helps in making it more effective at guessing a person's objectives and predicting intentions.

6. Results, Data Analysis, Comparison, and Discussion

To enable robots to adaptability and efficiently interact with people, it is necessary to apply ML techniques, allowing robots to perform target tasks and operate autonomously without direct human intervention while learning from humans. This mix of autonomy and adaptability learning from people is especially important for making robots, or any an autonomous system, work smoothly with humans in different places. To analyze data from different environments and to produce high-level information, the main research studies about HRI mentioned in the previous section are reviewed. These research studies use unsupervised and supervised learning methods. From a broader standpoint, it is worth emphasizing that most of these studies focus on approximating the basic interaction between humans and these intelligent machines (i.e. robots). For this well-defined motivation, Table IV provides a summary of these reviewed studies, where it is evident that the Neural learning algorithm and Reinforcement algorithm have 37% and 23% of the occurrences, respectively.

TABLE IV. STUDIES COMPARISON IN HUMAN-ROBOT INTERACTION (HRI)

Serial	Ref.	Year	Human and Robot Interaction	Machine Learning Algorithm
1	[46]	2024	Proximate	Neural
2	[47]	2023	Proximate	Deep
3	[49]	2023	Proximate	Continual
4	[54]	2023	Remotely	Deep
5	[38]	2023	Proximate	Reinforcement
6	[34]	2022	Remotely	SVM
7	[24]	2022	Proximate	Deep
8	[52]	2021	Proximate	Digital Twins
9	[53]	2021	Proximate	Deep
10	[55]	2020	Proximate	POMDP
11	[39]	2020	Proximate	Reinforcement
12	[56]	2020	Proximate	Neural
13	[57]	2020	Proximate	Neural
14	[58]	2020	Proximate	Neural
15	[59]	2020	Proximate	Neural
16	[40]	2020	Proximate	Reinforcement
17	[60]	2020	Proximate	Deep
18	[42]	2020	Proximate	Reinforcement
19	[43]	2020	Proximate	Reinforcement
20	[44]	2020	Proximate	Neural
21	[61]	2020	Proximate	Imitation
22	[62]	2020	Proximate	Deep
23	[63]	2020	Proximate	Neural
24	[32]	2020	Proximate	Clustering and Bayesian
25	[64]	2020	Proximate	Neural
26	[65]	2020	Proximate	Neural
27	[66]	2019	Proximate	Reinforcement
28	[67]	2019	Proximate	Neural
29	[41]	2019	Proximate	Reinforcement
30	[68]	2019	Proximate	Neural

According to the statistical analysis of this table (i.e. Table IV), it is observed that the identified occurrences in the investigation are present in 20% of research studies in HRI use Deep learning methods. Additionally, about 3% employ Continual, SVM, Digital Twins, POMDP, Imitation and Clustering and Bayesian learning methods. These findings pave the way for creating robots that are easier to understand and respond better in many different situations.

Additionally, 28 of the 30 research studies listed in the table (i.e., 93.33%) focus on proximate interactions, meaning that direct and/or close physical interaction between humans and robots is typically required in the context of HRI. This put emphasis on the importance of designing robots that can function effectively in close collaboration with humans where they share the same space; such as in personal assistance, education, and healthcare.

Figure 5 ensures that the total number of HRI research studies involved in using machine learning algorithms was 30. Among them, 11 research papers concentrated on using Artificial Neural Network (ANN) algorithms, while 7 and 6 research papers interested in exploring for both Reinforcement and Deep learning methods, respectively. However, in the reviewed studies, Continual, SVM, Digital Twins, POMDP, Imitation and Clustering and Bayesian algorithms were each employed in a single research study, respectively

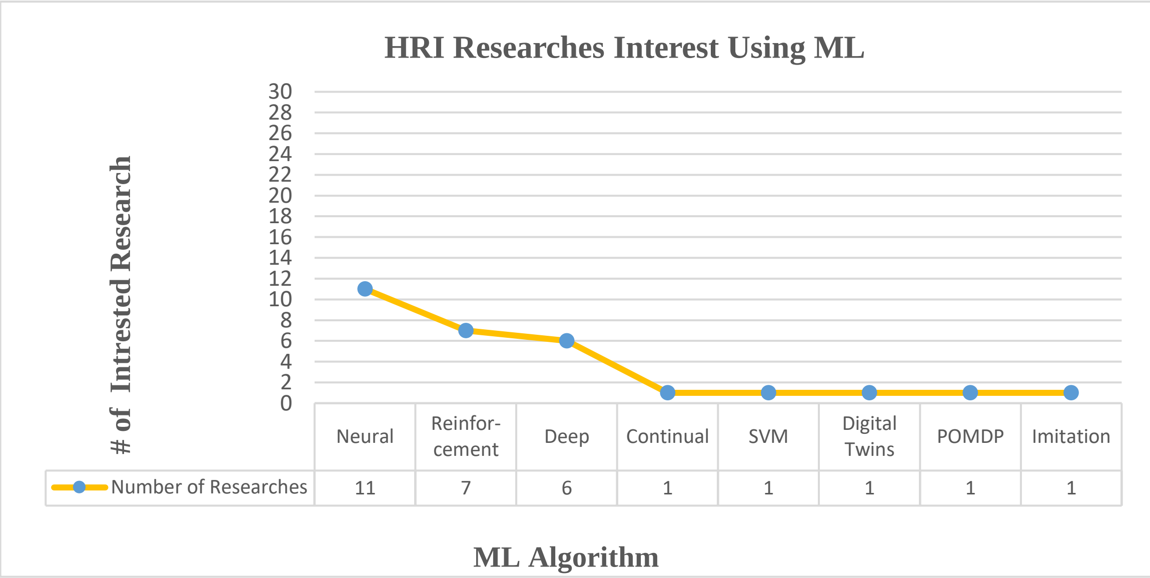

Fig.5: HRI Research Interest ML

As represented in the comparative breakdown of Figure 5, it is relatively clear that these significant results show that most recent studies that interest in HRI are focused on the Neural networks, Deep and Reinforcement machine learning algorithms on a physical application and approximate interaction. These outcomes demonstrate that neural networks and deep learning are crucial in enhancing how robots interact with humans, enabling robots to become smarter than before and more capable of understanding and responding to their human partners more effectively.

Again, these findings give sufficient sound evidence that there are many challenges with using HRI and machine learning algorithms, which makes it harder to use other methods like SVM. These challenges, in various ways, limit the use of these algorithms in social and mobile applications. Furthermore, this research demonstrates that despite the complexities involved in robot design, a robot's intelligence and performance can be significantly enhanced based on the learning methods used.

From another point of view, Table V reveals that the majority of interactions in the 30 research papers referenced in this table, accounting for 84%, were conducted through physical applications. On the other hand, social applications account for 13%, and mobile phone applications make up 3%, with a particular emphasis on tablet devices that facilitate user interaction through graphical buttons displaying information about device usage processes. It's important to note that voice control is becoming popular as an easy-to-use way to interact. So, people can use their voices to give orders to control devices, making it easier and more convenient to operate robots without needing to use physical controls. In other significant cases, visual sensors have been present in most research studies on HRI through machine learning to gather detailed and complex visual data that will later be analyzed to enhance the effectiveness of robotic systems.

TABLE V. TYPES OF SENSORS USED IN HRI APPLICATIONS

Serial	Ref.	Application	Data Sensing
1	[46]	Physical	Visual
2	[47]	Physical	Visual/Touch
3	[49]	Mobile	Visual
4	[54]	Physical	Touch
5	[38]	Physical	Visual/Voice
6	[34]	Physical	Visual
7	[24]	Social	Visual
8	[52]	Physical	Visual
9	[53]	Physical	Visual
10	[55]	Physical	Visual/Touch
11	[39]	Physical	Touch
12	[56]	Social	Voice
13	[57]	Physical	Motion/Touch
14	[58]	Physical	Motion/Touch
15	[59]	Physical	Touch
16	[40]	Physical	Touch
17	[60]	Physical	Visual
18	[42]	Physical	Force/Torque
19	[43]	Physical	Torque
20	[44]	Physical	Torque
21	[61]	Physical	Force/Position
22	[62]	Physical	Force/Torque/Position
23	[63]	Physical	Visual
24	[32]	Social	Visual
25	[64]	Social	Visual
26	[65]	Physical	Visual
27	[66]	Physical	Force/Touch
28	[67]	Physical	Force/Position/Velocity
29	[41]	Physical	Force/Torque/Motion
30	[68]	Physical	Visual/Force/Motion

7. HRI using ML Challenges and Potential Solutions

Humans use natural languages to accurately express their emotions and feelings whereas robots employ machine learning algorithms and Artificial Intelligence (AI) [1][2][17] Thus, natural language remains important for human-robot [2][17].To this aim, the interaction between a human operator and a robot needs to handle complex, variable, non-deterministic, and partially unknown environments. To deal with complex situations that may be difficult to interpret and understand, it is essential to exist communication, acceptable social responses, and good knowledge [1][2][17]. Because various artificial intelligence techniques are essential and necessary to allow the robot to understand and express feelings as part of the process interaction between humans and robots [66][1]. Thus, one of the biggest challenges with HRI is designing robots that can understand and respond to human emotions [69][1]. So, there is a need to understand the behavior of a human in interpreting and solving day-to-day problems and trying to come up with new algorithms that imitate humans[2][17].Furthermore, Nature Inspired Computing (NIC) refers to the fusion of nature, by itself, and AI to address and find practical solutions for the multifaceted problems that related to collecting, analyzing, and interpreting information related to the nature[17].Overall, one suggested practical solution to get around this problem is using more advanced sensors which can interpret body language, facial expressions, and the tone of the human voice.

Moreover, designing smart robots to work in unpredictable settings and environments such as public and large locations is another crucial challenge. Humans are characterized by flexibility and adapting to any environment, but robots are not. To overcome this problem a feasible and reliable solution is using more complex and advanced machine learning algorithms that enable robots to learn and adapt smartly to new settings.

One of the biggest and most important challenges in designing robotics is teaching them how to understand and imitate the people acts by themselves without needing instructions from human workers. This is very hard because human behavior is very complicated and changes a lot over time. Robots, as intelligent machines, in this way need to learn not just basic actions, but also the small, detailed ways people interact with each other. This, in a way or another, is vital for making robots cooperate better with humans.[1][2][19]

Despite robots cover many aspects, visible or invisible, of our life, they may be unacceptable by many humans in social roles due to their fear of losing their jobs, the privacy of their information, and ethical considerations, in addition to ensuring the safety of interaction between humans and robots and establishing trust between them [1][69][70]. This serious challenge, make a mean of unconfidently and, therefore, rated as some of the most extreme challenges which become unquestionably challenges to meet [1][69][70].

Aside from the above-mentioned discussion that has been presented in this research paper, most HRI systems that using ML algorithms are alike on average in terms of their performance where an algorithm may be the best choice for some types of problems, but at the same time, it may become to be a less suitable option for other kinds of problems. Then again, since most real-world problems have different needs and requirements depending on the industry and type of service, there isn't a practical one-size-fits-all algorithm that works well for every situation. This makes it challenging to find the true algorithm that adequately fits these specific needs [17].

8. Conclusion

Evidently, Human-to-robot interaction (HRI) has recently garnered a lot of attention in the academic community, in labs, in technology companies, and through the media. Given this interest, it is advisable to present a survey of studies in HRI that use machine learning (ML) to serve as an educational program for those outside the field and to encourage the community of academic and industrial researchers to initiate discussion of a unified view of HRI in this field. The main objective of this research study is to provide a review of the latest technology that includes most of the HRI studies using ML. Thus, a total of 30 research papers devoted to HRI were surveyed, evaluated, and analyzed to give the most ML algorithms implemented in the field of HRI. The comparison was made according to a set of factors including the types of robot sensors, HRI application and ML algorithm name, HRI type and algorithm types that is implemented in HRI type that is implemented in HRI. The reported results show that most recent studies that interest in HRI are focused on the supervised ML algorithms, including Neural Network and Reinforcement algorithms, and they can be applied in real-world tasks where robots interact physically with human workers. Besides helping to fill gaps in the literature, these added-value results may serve as a ground for further research in robotics development and AI enhancement.

Acknowledgments

The authors offer their heartfelt thanks to the editor and the esteemed anonymous reviewers of IJASCA for their insightful comments, valuable recommendations, and dedicated efforts that contributed to the improvement of this paper.

References

[1] S. O. Al-Hamouz, N. K. T. El-Omari, and A. M. Al-Naimat, "An ISO Compliant Safety System for Human Workers in Human-Robot Interaction Work Environment," in *2019 12th International Conference on Developments in eSystems Engineering (DeSE)*, Oct. 2019, vol. 20, no. 1, pp. 9–14, doi: 10.1109/DeSE.2019.00012.

[2] H. Owaied, N. K. T. El-Omari, and M. Abu-Arr'a, "User Interface for Automation of a Product Design Processes," *Ubiquitous Comput. Commun. J.*, vol. 4, no. 5, pp. 50–54, 2009.

[3] N. K. T. El-Omari, "Cloud IoT as a Crucial Enabler : a Survey and Taxonomy," *Mod. Appl. Sci.*, vol. 13, no. 2019, pp. 86–149, 2019, doi: 10.5539/mas.v13n8p86.

[4] V. O. Nyangaresi, N. K. T. El-Omari, and J. N. Nyakina, "Efficient Feature Selection and ML Algorithm for Accurate Diagnostics," *J. Comput. Sci. Res.*, vol. 4, no. 1, pp. 10–19, 2022, doi: 10.30564/jcsr.v4i1.3852.

[5] C. Bartneck, T. Belpaeme, F. Eyssel, T. Kanda, M. Keijsers, and S. Sabanovic, *Human–Robot Interaction: An Introduction*, vol. 38, no. 3. Cambridge University Press, 2022.

[6] J. Patel, P. Sonar, and C. Pinciroli, "On multi-human multi-robot remote interaction: a study of transparency, inter-human communication, and information loss in remote interaction," *Swarm Intell.*, vol. 16, no. 2, pp. 107–142, 2022, doi: 10.1007/s11721-021-00209-2.

[7] J. Mainprice, M. Gharbi, T. Simeon, and R. Alami, "Sharing Effort in Planning Human-Robot Handover Tasks," *Proc. - IEEE Int. Work. Robot Hum. Interact. Commun.*, no. 2012, pp. 764–770, 2012, doi: 10.1109/ROMAN.2012.6343844.

[8] P. Gustavsson, M. Holm, A. Syberfeldt, and L. Wang, "Human-robot collaboration -

Towards new metrics for selection of communication technologies," *Procedia CIRP 72*, vol. 72, no. March, pp. 123–128, 2018, doi: 10.1016/j.procir.2018.03.156.

[9] A. Bonarini, "Communication in Human-Robot Interaction," *Curr. Robot. Reports*, vol. 1, no. 4, pp. 279–285, 2020, doi: 10.1007/s43154-020-00026-1.

[10] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: A survey," *Foundations and Trends in Human-Computer Interaction*, vol. 1, no. 3. pp. 203–275, 2007, doi: 10.1561/1100000005.

[11] S. Ikemoto, H. Ben Amor, T. Minato, B. Jung, and H. Ishiguro, "Physical human-robot interaction: Mutual learning and adaptation," *IEEE Robot. Autom. Mag.*, vol. 19, no. 4, pp. 24–35, 2012, doi: 10.1109/MRA.2011.2181676.

[12] S. E. Spisak and B. Indurkhya, "A Study on Social Exclusion in Human-Robot Interaction," *Electron.*, vol. 12, no. 7, pp. 1–17, 2023, doi: 10.3390/electronics12071585.

[13] Y. Chen *et al.*, "Human mobile robot interaction in the retail environment," *Sci. Data*, vol. 9, no. 1, pp. 1–10, 2022, doi: 10.1038/s41597-022-01802-8.

[14] A. Bonarini, F. Garzotto, M. Gelsomini, M. Romero, F. Clasadonte, and A. N. C. Yilmaz, "A huggable, mobile robot for developmental disorder interventions in a multi-modal interaction space," *25th IEEE Int. Symp. Robot Hum. Interact. Commun. RO-MAN 2016*, pp. 823–830, 2016, doi: 10.1109/ROMAN.2016.7745214.

[15] R. S. Andersen, O. Madsen, T. B. Moeslund, and H. Ben Amor, "Projecting robot intentions into human environments," *25th IEEE Int. Symp. Robot Hum. Interact. Commun. RO-MAN 2016*, pp. 294–301, 2016, doi: 10.1109/ROMAN.2016.7745145.

[16] B. Zughoul, N. K. T. El-Omari, and M. Al-Refai, "Using deep learning methods in detecting the critical success factors on the implementation of cloud ERP," *Int. J. Bus. Inf. Syst.*, vol. 44, no. 2, pp. 219–248, 2023, doi: 10.1504/ijbis.2020.10036721.

[17] N. K. T. El-Omari, "Sea Lion Optimization Algorithm for Solving the Maximum Flow Problem," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 20, no. 8, p. 30, 2020, doi: 10.22937/IJCSNS.2020.20.08.5.

[18] N. K. T. El-Omari, "A Hybrid Approach for Segmentation and Compression of Compound Images," 2008.

[19] N. T. V. Tuyen, A. L. Georgescu, I. Di Giulio, and O. Celiktutan, "A Multimodal Dataset for Robot Learning to Imitate Social Human-Human Interaction," in *ACM/IEEE International Conference on Human-Robot Interaction*, 2023, vol. 1, no. 1, pp. 238–242, doi: 10.1145/3568294.3580080.

[20] N. K. T. El-Omari, "An Efficient Two-Level Dictionary-Based Technique for Segmentation and Compression Compound Images," *Mod. Appl. Sci.*, vol. 14, no. 4, p. 52, 2020, doi: 10.5539/mas.v14n4p52.

[21] R. N. Albustanji, S. Elmanaseer, and A. A. A. Alkhatib, "Robotics: Five Senses plus One—An Overview," *Robot. J.*, vol. 12, no. 3, pp. 1–23, 2023, doi: 10.3390/robotics12030068.

[22] M. Andres Zamora Hernandez, E. Caldwell Marin, J. Garcia-Rodriguez, J. Azorin-Lopez, and M. Cazorla, "Automatic learning improves human-robot interaction in productive environments: A review," in *Computer Vision: Concepts, Methodologies, Tools, and Applications*, vol. 7, no. 3, 2018, pp. 2014–2024.

[23] S. Suzuki, Y. Mitsukura, H. Takimoto, T. Tanabata, N. Kimura, and T. Moriya, "A human tracking mobile-robot with face detection," *IECON Proc. (Industrial Electron. Conf.*, no. December, pp. 4217–4222, 2009, doi: 10.1109/IECON.2009.5415074.

[24] N. Ottakath *et al.*, "ViDMASK dataset for face mask detection with social distance

measurement," *Displays*, vol. 73, no. January, pp. 1–13, 2022, doi: 10.1016/j.displa.2022.102235.

[25] M. I. Adawy, S. A. Nor, and M. Mahmuddin, "Data redundancy reduction in wireless sensor network," *J. Telecommun. Electron. Comput. Eng.*, vol. 10, no. 1–11, pp. 1–6, 2018.

[26] H. A. Pierson and M. S. Gashler, "Deep learning in robotics: a review of recent research," *Adv. Robot.*, vol. 31, no. 16, pp. 821–835, 2017, doi: 10.1080/01691864.2017.1365009.

[27] W. Fung and Y. Liu, "Feature extraction of robot sensor data using factor analysis for behavior learning," *J. Adv. Comput. Intell. Intell. Informatics*, vol. 8, no. 3, pp. 284–294, 2004, doi: 10.20965/jaciii.2004.p0284.

[28] R. Gopalapillai, J. Vidhya, D. Gupta, and T. S. B. Sudarshan, "Classification of robotic data using artificial neural network," *2013 IEEE Recent Adv. Intell. Comput. Syst. RAICS 2013*, pp. 333–337, 2013, doi: 10.1109/RAICS.2013.6745497.

[29] M. I. Adawy, M. Tahboush, O. Aloqaily, and W. Abdulraheem, "Man-In-The Middle Attack Detection Scheme on Data Aggregation in Wireless Sensor Networks," *Int. J. Adv. Soft Comput. its Appl.*, vol. 15, no. 2, pp. 179–193, 2023, doi: 10.15849/IJASCA.230720.12.

[30] M. Andres Zamora Hernandez, E. Caldwell Marin, J. Garcia-Rodriguez, J. Azorin-Lopez, and M. Cazorla, "Automatic learning improves human-robot interaction in productive environments: A review," in *Computer Vision: Concepts, Methodologies, Tools, and Applications*, vol. 2, 2018, pp. 2014–2024.

[31] M. Tahboush, M. Agoyi, and A. Esaid, "Multistage security detection in mobile ad-hoc network (MANET)," *Int. J. Eng. Trends Technol.*, vol. 68, no. 11, pp. 97–104, 2020, doi: 10.14445/22315381/IJETT-V68I11P213.

[32] S. Vinanzi, A. Cangelosi, and C. Goerick, "The Role of Social Cues for Goal Disambiguation in Human-Robot Cooperation," *29th IEEE Int. Conf. Robot Hum. Interact. Commun. RO-MAN 2020*, pp. 971–977, 2020, doi: 10.1109/RO-MAN47096.2020.9223546.

[33] A. Moubayed, M. Injadat, A. B. Nassif, H. Lutfiyya, and A. Shami, "E-Learning: Challenges and Research Opportunities Using Machine Learning Data Analytics," *IEEE Access*, vol. 6, no. 2018, pp. 39117–39138, 2018, doi: 10.1109/ACCESS.2018.2851790.

[34] J. Wang, M. R. Pradhan, and N. Gunasekaran, "Machine learning-based human-robot interaction in ITS," *Inf. Process. Manag.*, vol. 59, no. 1, pp. 1–11, 2022, doi: 10.1016/j.ipm.2021.102750.

[35] N. K. T. El-Omari, A. H. Omari, O. F. Al-badarneh, and H. Abdel-jaber, "Scanned Document Image Segmentation Using Back-Propagation Artificial Neural Network Based Technique," *Int. J. Comput. Commun. Control*, vol. 6, no. 4, pp. 183–190, 2012.

[36] R. A. Mouha, "Deep Learning for Robotics," *J. Data Anal. Inf. Process.*, vol. 9, no. 2, pp. 63–76, 2021, doi: 10.4236/jdaip.2021.92005.

[37] U. Kartoun, H. Stern, and Y. Edan, "A Human-Robot Collaborative Reinforcement Learning Algorithm," *J. Intell. Robot. Syst.*, no. April, 2010, doi: 10.1007/s10846-010-9422-y.

[38] F. Munguia-Galeano, S. Veeramani, J. D. Hernandez, Q. Wen, and Z. Ji, "Affordance-Based Human-Robot Interaction With Reinforcement Learning," *IEEE Access*, vol. 11, no. March, pp. 31282–31292, 2023, doi: 10.1109/ACCESS.2023.3262450.

[39] S. C. Akkaladevi, M. Plasch, A. Pichler, and M. Ikeda, "Towards reinforcement based learning of an assembly process for human robot collaboration," *Procedia Manuf.*, vol. 38, no. Faim 2019, pp. 1491–1498, 2020, doi: 10.1016/j.promfg.2020.01.138.

[40] A. Ghadirzadeh, X. Chen, W. Yin, Z. Yi, M. Bjorkman, and D. Kragic, "Human-Centered Collaborative Robots with Deep Reinforcement Learning," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 566–571, 2021, doi: 10.1109/LRA.2020.3047730.

[41] A. Tabrez, S. Agrawal, and B. Hayes, "Explanation-Based Reward Coaching to Improve Human Performance via Reinforcement Learning," *ACM/IEEE Int. Conf. Human-Robot Interact.*, vol. 2019-March, pp. 249–257, 2019, doi: 10.1109/HRI.2019.8673104.

[42] W. Lu, Z. Hu, and J. Pan, "Human-Robot Collaboration using Variable Admittance Control and Human Intention Prediction," *IEEE Int. Conf. Autom. Sci. Eng.*, vol. 2020-Augus, pp. 1116–1121, 2020, doi: 10.1109/CASE48305.2020.9217040.

[43] L. Roveda *et al.*, "Model-Based Reinforcement Learning Variable Impedance Control for Human-Robot Collaboration," *J. Intell. Robot. Syst. Theory Appl.*, vol. 100, no. 2, pp. 417–433, 2020, doi: 10.1007/s10846-020-01183-3.

[44] M. Wu, Y. He, and S. Liu, "Adaptive impedance control based on reinforcement learning in a human-robot collaboration task with human reference estimation," *Int. J. Mech. Control*, vol. 21, no. 1, pp. 21–31, 2020.

[45] A. M. Andrew, *The Handbook of Brain Theory and Neural Networks*, vol. 28, no. 9. 1999.

[46] J. Patalas-Maliszewska, A. Dudek, G. Pajak, and I. Pajak, "Working toward Solving Safety Issues in Human–Robot Collaboration: A Case Study for Recognising Collisions Using Machine Learning Algorithms," *Electron.*, vol. 13, no. 4, pp. 1–16, 2024, doi: 10.3390/electronics13040731.

[47] M. Rezayati, G. Zanni, Y. Zaoshi, D. Scaramuzza, and H. W. Van De Venn, "Improving safety in physical human-robot collaboration via deep metric learning," *IEEE Int. Conf. Emerg. Technol. Fact. Autom. ETFA*, vol. 2022-Septe, no. September 2022, 2023, doi: 10.1109/ETFA52439.2022.9921623.

[48] S. Li, "Introduction and Perspective of Deep Metric Learning," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4320050.

[49] A. Ayub, Z. De Francesco, P. Holthaus, C. L. Nehaniv, and K. Dautenhahn, "Continual Learning through Human-Robot Interaction - Human Perceptions of a Continual Learning Robot in Repeated Interactions," *arXiv.org*, pp. 1–36, 2023, doi: 10.48550/arXiv.2305.16332.

[50] N. S. Simul, N. M. Ara, and M. S. Islam, "A support vector machine approach for real time vision based human robot interaction," in *19th International Conference on Computer and Information Technology, ICCIT 2016*, 2017, pp. 496–500, doi: 10.1109/ICCITECHN.2016.7860248.

[51] N. Briquet-Kerestedjian, A. Wahrburg, M. Grossard, M. Makarov, and P. Rodriguez-Ayerbe, "Using neural networks for classifying human-robot contact situations," *2019 18th Eur. Control Conf. ECC 2019*, pp. 3279–3285, 2019, doi: 10.23919/ECC.2019.8795649.

[52] A. A. Malik and A. Brem, "Digital twins for collaborative robots: A case study in human-robot interaction," *Robot. Comput. Integr. Manuf.*, vol. 68, p. 102092, Apr. 2021, doi: 10.1016/j.rcim.2020.102092.

[53] B. Yan, P. Fan, X. Lei, Z. Liu, and F. Yang, "A real-time apple targets detection method for picking robot based on improved YOLOv5," *Remote Sens.*, vol. 13, no. 9, pp. 1–23, 2021, doi: 10.3390/rs13091619.

[54] M. BARSTUĞAN and Z. OSMANPAŞAOĞLU, "Deep Learning Based Human Robot Interaction With 5G Communication," *Konya J. Eng. Sci.*, vol. 11, no. 2, pp. 423–438, 2023, doi: 10.36306/konjes.1228275.

[55] M. A. Ahmad, M. Ourak, C. Gruijthuijsen, J. Deprest, T. Vercauteren, and E. Vander

Poorten, "Deep learning-based monocular placental pose estimation: towards collaborative robotics in fetoscopy," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 9, pp. 1561–1571, 2020, doi: 10.1007/s11548-020-02166-3.

[56] M. Chen, S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa, "Trust-Aware Decision Making for Human-Robot Collaboration," *ACM Trans. Human-Robot Interact.*, vol. 9, no. 2, pp. 1–23, Jun. 2020, doi: 10.1145/3359616.

[57] X. Chen, Y. Jiang, and C. Yang, "Stiffness Estimation and Intention Detection for Human-Robot Collaboration," *Proc. 15th IEEE Conf. Ind. Electron. Appl. ICIEA 2020*, pp. 1802–1807, 2020, doi: 10.1109/ICIEA48937.2020.9248186.

[58] X. Chen, N. Wang, H. Cheng, and C. Yang, "Neural Learning Enhanced Variable Admittance Control for Human-Robot Collaboration," *IEEE Access*, vol. 8, pp. 25727–25737, 2020, doi: 10.1109/ACCESS.2020.2969085.

[59] A. Cunha *et al.*, "Towards collaborative robots as intelligent co-workers in human-robot joint tasks: What to do and who does it?," *52nd Int. Symp. Robot. ISR 2020*, pp. 141–148, 2020.

[60] T. H. S. Li, P. H. Kuo, T. N. Tsai, and P. C. Luan, "CNN and LSTM Based Facial Expression Analysis Model for a Humanoid Robot," *IEEE Access*, vol. 7, pp. 93998–94011, 2020, doi: 10.1109/ACCESS.2019.2928364.

[61] A. Sasagawa, K. Fujimoto, S. Sakaino, and T. Tsuji, "Imitation learning based on bilateral control for human-robot cooperation," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 6169–6176, 2020, doi: 10.1109/LRA.2020.3011353.

[62] J. Zhang, H. Liu, Q. Chang, L. Wang, and R. X. Gao, "Recurrent neural network for motion trajectory prediction in human-robot collaborative assembly," *CIRP Ann.*, vol. 69, no. 1, pp. 9–12, 2020, doi: 10.1016/j.cirp.2020.04.077.

[63] W. Wojtak, F. Ferreira, P. Vicente, L. Louro, E. Bicho, and W. Erlhagen, "A neural integrator model for planning and value-based decision making of a robotics assistant," *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3737–3756, 2021, doi: 10.1007/s00521-020-05224-8.

[64] L. Van Der Spaa, M. Gienger, T. Bates, and J. Kober, "Predicting and Optimizing Ergonomics in Physical Human-Robot Cooperation Tasks," *Proc. - IEEE Int. Conf. Robot. Autom.*, no. Icra, pp. 1799–1805, 2020, doi: 10.1109/ICRA40945.2020.9197296.

[65] S. Murata, W. Masuda, J. Chen, H. Arie, T. Ogata, and S. Sugano, "Achieving human--robot collaboration with dynamic goal inference by gradient descent," in *Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12--15, 2019, Proceedings, Part II 26*, 2019, pp. 579–590, doi: 10.1007/978-3-030-36711-4_49.

[66] L. Peternel, C. Fang, N. Tsagarakis, and A. Ajoudani, "A selective muscle fatigue management approach to ergonomic human-robot co-manipulation," *Robot. Comput. Integr. Manuf.*, vol. 58, pp. 69–79, 2019, doi: 10.1016/j.rcim.2019.01.013.

[67] L. Roveda, S. Haghshenas, M. Caimmi, N. Pedrocchi, and L. M. Tosatti, "Assisting operators in heavy industrial tasks: On the design of an optimized cooperative impedance fuzzy-controller with embedded safety rules," *Front. Robot. AI*, vol. 6, no. August, 2019, doi: 10.3389/frobt.2019.00075.

[68] L. Yan, X. Gao, X. Zhang, and S. Chang, "Human-robot collaboration by intention recognition using deep LSTM neural network," in *Proceedings of the 8th International Conference on Fluid Power and Mechatronics, FPM 2019*, 2019, pp. 1390–1396, doi: 10.1109/FPM45753.2019.9035907.

[69] O. C. Chikwendu *et al.*, "Human-Robot Interaction Enhancement Through Ergonomics

and Human Factors : Future Directions," *Int. J. Eng. Res. Dev.*, vol. 19, no. 6, p. .34-40, 2023.

[70] M. Tahboush, M. Adawy, and O. Aloqaily, "PEO-AODV: Preserving Energy Optimization Based on Modified AODV Routing Protocol for MANET," *Int. J. Adv. Soft Comput. its Appl.*, vol. 15, no. 2, pp. 263–277, 2023, doi: 10.15849/IJASCA.230720.18.

Mohammad I. Adawy is Assistant Professor at the Department of Networks System and Security. The World Islamic Sciences and Education University, Amman, Jordan. His main teaching and research interests include human-robot interaction, research methodology and results analysis. He has published several research articles in international journals of computer networks and

security.Finally, the author can be contacted atMohammad.adawi@wise.edu.jo

Nidhal K. T. El-Omari received his B.Sc. in Computer Science and his M. Eng. degree in Computer Engineering in 1986 and 2005, respectively, both from Yarmouk University, Irbid- Jordan. In 1989, he received his Higher Diploma of Branch Automation Officer from the Department of Defense (DoD), Fort Gordon / Georgia-USA. In 2008, he received a doctorate in Artificial Intelligence and Digital Image Processing & Encryption from The Arab Academy for Banking and Financial Science (AABFS).

Prof. El-Omari joined the Information Technology Directorate of the Jordanian Ministry of Defense in 1986 and retired in 2009. During those 24 years, he chaired a number of IT-related departments including the Systems Follow-up Department, Technical Support Department, and Automation Department. He has been at the Faculty of IT since 2009, WISE University, during which he worked as the director of the Computer Center, the Chair of the Department of Computer Science and Basic Science, and the head of the Department of Software Engineering. Currently, he is the Dean of the College of Information Technology. Since 2021, he has been awarded the title of Full Professor. His main teaching and research interests include, but are not limited to, image compression & segmentation, evolutionary computation, heuristic optimization, and methodologies for building both secure and strategical-efficient software. Prof. El-Omari has authored/co-authored two computer books and published more than thirty research articles and conference papers in top-quality journals and conference proceedings.

Hasan Abualese is currently an assistant professor of software engineering at Al-Ahliyya Amman University, Jordan. His research activities focus on software testing and quality assurance, web security and human-robot interaction. He can be contacted at h.abualese@ammanu.edu.jo

Abdulwadood Alawadhi is an assistant professor of Computer Networks at Universiti Utara Malaysia (UUM), Malaysia. His research activities focus on computer networks, wireless sensor networks, and human-robot interaction. He can be contacted at alawadhi@uum.edu.my