Web Perspectives in Robotics Applications: Commonsense Knowledge, Autonomous Vehicles and Human-Robot Collaboration

CHRISTOPHER J. CONTI, APARNA S. VARDE and WEITIAN WANG
Department of Computer Science, Montclair State University, NJ, USA

The realms of commonsense knowledge and reasoning, vehicle automation with full as well as partial autonomy, and human-robot collaboration, present growing areas of research in recent times, with much of the concerned data being disseminated through the Web and devices encompassing IoT (Internet of Things); the data per se being heterogeneous including plain text, images, audiovisuals, hypertext and hypermedia. Due to the advent of autonomous vehicles, there is a greater need for the embodiment of commonsense knowledge within their development in order to simulate subtle, intuitive aspects of human judgment. The field of robotics has often encountered collaborative tasks between humans and robots to enhance the respective activities involved and produce better results than humans or robots would achieve working by themselves. Accordingly, this article outlines and organizes some of the research occurring in these areas along with its Web perspectives and applications. Context related to human-robot collaboration and commonsense knowledge appears via a survey of the literature. Vehicle automation is significant with the relevant studies: its definition and methods of improvement are of focus in the article. Some work in this area makes an impact on smart manufacturing. There is discussion on how human-robot collaboration is beneficial, and how commonsense knowledge is useful for the collaboration to occur in an enhanced manner. This article would be potentially interesting to various communities, e.g. AI professionals, Web developers, robotics engineers, and data scientists.

DOI: 10.1145/3583849.3583853 http://doi.acm.org/10.1145/3583849.3583853

1. INTRODUCTION

Human-robot collaboration (HRC) is useful in helping companies and human workers; however, robots need several skills in order to work with humans effectively. Collaborative robots, sometimes referred to as "cobots", need to learn from users easily and utilize planning that can handle tasks in dynamic real-world situations [20]. An area of research relating to human-robot collaboration is commonsense knowledge. The term commonsense knowledge (CSK) involves a basic, intuitive understanding of real-world objects, their properties and associations and the manner in which interact with each other. CSK is inherent in humans and is typically too subtle for machines such as robots to comprehend, yet it is necessary for robots in handling everyday real-world situations in a manner analogous to humans [12, 43]. Reasoning based on CSK is important to enhance planning and decision-making. Robots also need to detect objects in images in order to determine how to act in their working environment, which brings us to the field of object recognition. This can benefit from CSK as well. Automated vehicles including self-driving cars, trains etc. constitute a growing field and can be helpful for people who are unable to drive, e.g. the elderly or the disabled, and for those who prefer automation in driving for saving time and effort. Automation in driving can be cost-effective and efficient compared to human driving. In the realm of automated vehicles, we have autonomous vehicles that make independent well-informed decisions on their own without human intervention. Additionally, there are semi-autonomous vehicles that are capable of such decision-making but are typically intervened by humans. Such types of vehicles are useful in modern applications and there is tremendous research in these areas in order to achieve the vision of fully autonomous vehicles being commonplace in the near future. Much of the data pertaining to these areas is disseminated via the Web through desktops, laptops, tablets, smartphones and other Web-enabled devices. The Web has emerged as an indispensable resource to garner information and use it for a myriad of tasks, spanning commonsense knowledge, collaborative robotics and autonomous vehicles.

This article presents a technical survey that outlines relevant research in the aforementioned areas, e.g. how robots can supplement human workers, how commonsense knowledge can be useful here, how such research is beneficial to autonomous vehicles, and how it makes a broader impact on smart manufacturing, entailing the use of AI and robotics to enhance manufacturing processes. It asserts that data dissemination via the Web and Web-enabled devices is crucial in all these areas. Furthermore, simulations often precede real-world experiments, especially given the ever-growing trend in moving online. In particular, the paper mainly focuses on the following aspects, with respect to their Web perspectives and suitable applications.

- Commonsense Knowledge
- Autonomous Vehicles
- Human-Robot Collaboration

The organization of the rest of this article is thus into these respective sections, each focusing on one of the aforementioned aspects. We also explain within each section how the aspects relate to each other, especially with reference to their usefulness in real-world applications. In what follows, several related research works about commonsense knowledge, including the concerned Web-based repositories appear in Section 2. Autonomous vehicles are of focus in Section 3 with a mention of IoV, i.e. Internet of Vehicles, in this matter. We expound on human-robot collaboration in Section 4, with a view to smart manufacturing and Internet of Things, i.e., IoT. Finally, we put forth conclusions in Section 5 along with open issues for further research and development.

2. COMMONSENSE KNOWLEDGE

Commonsense knowledge (CSK) focuses on day-to-day entities, how they are related and how they interact [43]. It is the type of knowledge that is often too obvious to humans, e.g. the fact that the steering wheel of a car is always in the front and never behind. Machines are very competent at memorizing encyclopedic facts, e.g. prices, makes and models of cars, but have a much more difficult time with commonsense knowledge, especially in the absence of explicit training [42]. For instance, while machines may be better at memorizing weather patterns and their relevance to driving conditions, they may not be able to know that people enjoy walking when it is mildly cloudy rather than rainy or overly hot. A human would easily know this due to inherent common sense. While machines can identify a vehicle, humans would intuitively be able to tell if the vehicle is parked, pulling out from parking, or moving. Another example of commonsense knowledge is that a larger person can usually carry more weight than a smaller one. Additionally, if a person is walking in a crosswalk with their phone, they are unlikely to be paying careful attention to their surroundings. Hence, while machines are excellent at storing and using encyclopedic knowledge, often overshadowing humans in such tasks, they lag behind humans with respect to commonsense knowledge, something humans inherently possess. Formalizing such knowledge is currently an area where it would be useful for machines to improve, hence motivating the need for commonsense knowledge bases [42]. Hence, there are various commonsense knowledge bases developed that serve as useful sources of real-world knowledge, encompassing semantics and pragmatics. These are available on the Web for universal access, and can facilitate extraction of pertinent information via Web-enabled devices. This is highly beneficial to the field of robotics, since robots are machines that are trying to achieve near-human capabilities. The acquisition of commonsense knowledge in terms of its extraction as well as compilation constitutes systematic tasks entailing numerous challenges, and the resulting Web-based knowledge bases need to undergo intrinsic as well as extrinsic evaluation so that they can be beneficial to several applications such as text mining in specific domains [36]. Intrinsic evaluation pertains to how good the acquired knowledge is, and is measurable by standard metrics such as precision and recall, adapted with reference to context. Extrinsic evaluation relates to how useful the acquired knowledge is, and often needs more studies with reference to context along with the concerned criteria. This can occur in an application-specific manner, outlining evaluation criteria accordingly.

We claim that CSK can potentially help in various areas, such as comprehending texts, computer vision, planning and reasoning for various applications related to robotics, e.g. [2], [6]. Depending on the nature of the data: 2-dimensional and 3-dimensional images, audiovisual data, natural language text, voice and speech, hypertext, hypermedia etc. there is different types of CSK. For instance in natural language, simple programs by default handle individual words and short phrases in a passage without the full context of the passage, while CSK helps with finding enhanced results analogous to humans. Computer vision can improve through common sense by continuing to track position of objects after they leave the frame, which is highly helpful with images. Robots need reasoning to handle unexpected events, such as a catering robot not serving wine if the glass is broken. Utilizing commonsense knowledge therefore has a variety of benefits in real-world scenarios.

A widely-referenced commonsense knowledge base, WebChild [42], is an initiative of the Max Planck Institute for Informatics, Germany. WebChild is a huge repository of everyday commonsense concepts, properties and relationships extracted from the Web such that the focus is clearly more towards commonsense per se (as opposed to encyclopedic knowledge). Each concept in this commonsense knowledge base appears with respect to its corresponding real-world domain, its comparable concepts, its physical parts if applicable, its associated activities, its relevant properties and its typical locations. Suitable pictures of commonsense concepts occur as appropriate. Users can browse these concepts via the WebChild commonsense browser [42]. We present a relevant partial snapshot of this browser interface in Figure 1. This describes the concept of "vehicle" using a picture, an explanation with simple terms on its activities, physical parts, spatial proximity etc. Note that further clicking on "conveyance" in the "type-of" box here would lead to other details. Similarly, other boxes would provide more information upon clicking. Likewise, WebChild offers insights into several useful commonsense concepts and can be beneficial in conjunction with other systems in real-world applications, e.g. object recognition systems such as YOLO [38], to enhance their commonsense capabilities. There is a plethora of commonsense knowledge bases developed over the years, as discussed in recent research on CSK, e.g., a tutorial in ACM CIKM (Conference on Information and Knowledge Management) the contents of which are outlined in the ACM SIGMOD Record journal [43], and another tutorial at ACM WSDM (Web Search and Data Mining conference). We briefly mention a few research works here [7, 14, 40] that seem relevant to robotics, in particular considering their usefulness in automated vehicles, and HRC.

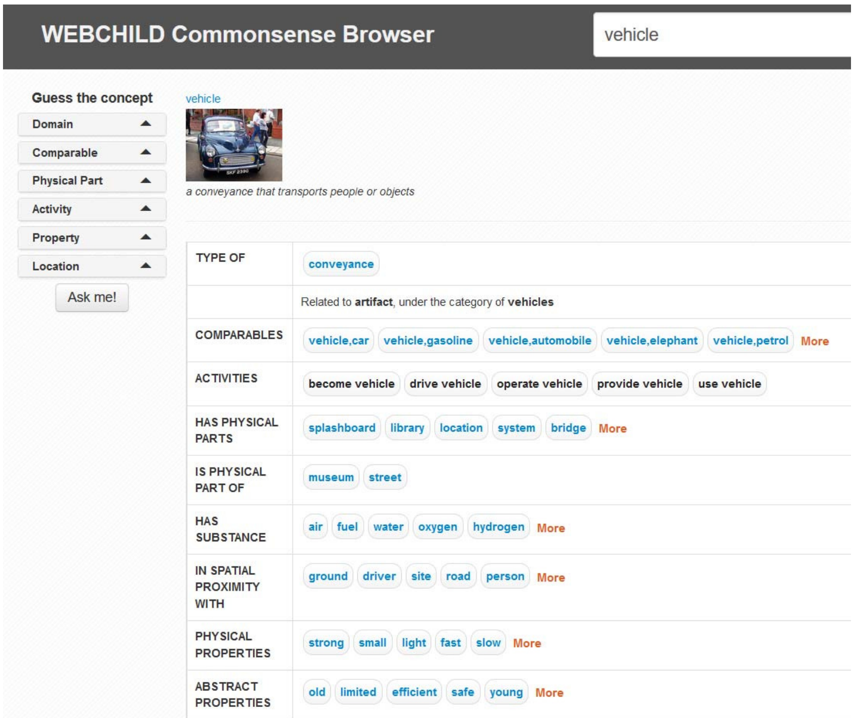

Fig. 1. Partial snapshot of the WebChild commonsense browser [42]

An integrated and challenging knowledge base called ARC [7], namely the Ai2 Reasoning Challenge, is the product of Allen Institute for Artificial Intelligence (Ai2), Seattle, WA. This system encompasses various knowledge types such as definition, basic facts and properties, structure, processes and causal, experiments, etc. An example of definition-based knowledge is, “worldwide increase in temperature is called global warming”. Additionally, it comprises various reasoning types including question logic, linguistic matching, comparison, spatial etc. For example, spatial reasoning can entail the following, “a sidewalk will feel hottest on a warm clear day in direct sunlight, as opposed to in the shade or under a picnic table”. While such reasoning seems rather implicit and intuitive to humans, it needs explicit programming into AI systems, which lack inherent commonsense premises. Such commonsense knowledge and reasoning can then be useful in understanding that pedestrians will be more prevalent in sidewalks in the shade than those exposed to direct sunlight on a warm clear day, and hence an automated vehicle can use such knowledge when it is nearing sidewalks. Dissemination of such information in a ubiquitous manner, e.g. through

Web-enabled devices with IoT, can be useful for reasoning in several applications related to robotics.

Winogrande [40] is a knowledge base developed as an enhancement to an earlier one called Winograd, both of these being in the category of commonsense reasoning, highly useful in communication. Winogrande deals with twin sentences, trying to disambiguate them with reference to context, providing adequate training data for AI systems to learn. For example, in the sentences: "The part does not fit in the socket because *it* is too large" versus, "The part does not fit in the socket because *it* is too small", the term "*it*" in the first sentence refers to the "part", while that in the second sentence refers to the "socket". This is too obvious to humans but needs explicit comprehension by AI systems, achieved through such knowledge bases. Such information is extremely useful for human-robot collaboration. For example, if a human and a robot collaborate in smart manufacturing, a human would see a part and socket and acquire such knowledge intuitively. Based on that, if the human communicates with the robot through a given sentence, the robot should be able to fathom the correct meaning with reference to context. Accordingly, Winogrande offers a dataset with around 44,000 problems with advanced scale and hardness to serve as a tough benchmark for commonsense reasoning. Consider the following sentence pairs that are harder than the previous example. "John wakes up at 9am while Tom wakes up at 6am, so he has less time to drive to work" versus, "John wakes up at 9am while Tom wakes up at 6am, so he has more time to drive to work". In these sentences, "he" in the first sentence refers to "John" while in the second one "he" refers to "Tom". While this may take a trifling longer than the previous example to fathom, it is still rather obvious to humans due to subtle commonsense reasoning, but is rather hard for an AI system. Hence, such knowledge bases, along with their Web access, can augment AI systems thus enabling them to learn better. The knowledge in the latter two sentences is useful with respect to autonomous driving, especially if voice-operated commands are involved for passengers and their commutes.

An interesting knowledge base called DoQ, i.e. Distribution over Quantities [14] addresses some quantitative aspects of commonsense knowledge, for example facts such as, "trucks are larger than cars". It considers various measurement types including length, mass, currency, temperature etc. It captures around 120 million unique tuples of the type (object, measurement) along with numerous instances of their real-world occurrences. They propound an unsupervised approach to garner quantitative information from vast quantities of web data, and deploy it to build a huge knowledge base consisting of distributions over physical quantities related to a multitude of objects. Such information is beneficial for AI systems, so they can function better. It is certainly useful in the context of object detection, such as distinguishing animals from each other in a given picture frame (e.g. a cat versus a cow). This in turn, can be applicable to autonomous vehicles as well as human-robot collaboration, both of which require precise object recognition. Additionally, such facts can be directly useful in planning tasks for human-robot collaboration, e.g. a fact such as "a car-seat is heavier than a rear-view mirror" can be useful in vehicle assembly where humans and robots work together. Likewise, the fact, "traffic lights typically stop for one minute" is directly useful in autonomous driving. Hence, such commonsense knowledge bases, others derived from them, and their respective enhancements can be extremely beneficial to autonomous driving and human-robot collaboration, considering the actual task of object detection as well as a related task of communication.

In addition to the knowledge bases described here, people frequently use translation tools as evidenced by the fact that Google Translate has over 500 million daily users. However, translation tools often have issues with collocations, i.e. the typical manner of using correct colloquial expressions embodying commonly used terms, such as "rat race" rather than "rat rush" or "clear sky" rather than "pure sky" [30]. Fixing odd collocations can help with querying the Internet for information and correcting errors in machine translated documents by indirectly harnessing commonsense knowledge via collocations. A tool called AwkChecker [30] gears towards assisting non-native speakers in correcting collocation errors. It flags such errors and recommends replacements that reflect common parlance. Another system called CollOrder is an example of a software tool for fixing odd collocations [46]. CollOrder conducts POS (part-of-speech) tagging and then searches its knowledge bases encompassing correct native speaker English (e.g. the British National Corpus: BNC) for appropriate collocations corresponding toe the same parts of speech. Suggestions from the knowledge bases are then ranked and filtered, and constitute the output ordered by frequency. Software systems such as AwkChecker and CollOrder can have practical applications from a CSK standpoint. Conveying the precise meaning of an expression as per common parlance is very important in AI systems, especially to avoid misunderstanding, and consequently prevent wrong outcomes due to miscommunication. Hence, such collocation fixing tools on the lines of AwkChecker and CollOrder can enhance communication in autonomous vehicles as well as human-robot collaboration. Making such tools available via IoT, e.g. on passengers' mobile devices can enhance user experiences in communication.

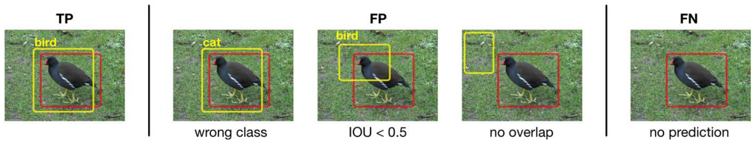

Fig. 2. Types of incorrect detections in object recognition, applicable to heterogeneous data [15].

A system that automatically predicts where object recognition would fail would be useful for training any object detection system via utilizing a hostile dataset. For instance, Figure 2 illustrates types of incorrect detections using a true positive as a correct example. Here, a true positive (TP) is a bird detected as a bird within the given bounding box, which is a correct classification. Errors or detection failures include: a bird detected as a cat that entails predicting the wrong class; a bird detected outside its correct bounding box, i.e. an example of a false positive (FP); and no overlap between the bounding boxes with no labeling of the bird at all. Finally, a false negative (FN) entails no prediction, i.e. absence of detection of the bird within the image even though there exists one. According to a recent study focusing on object detection [15], commonsense knowledge can be used to automatically predict such failures and hence help to provide benchmarks in object recognition. This leads to the development of a system known as CSK-SNIFFER [16], as synopsized in Figure 3, to automatically find where errors could potentially occur by using spatial commonsense, i.e. the knowledge of where two objects are usually located relative to each other, e.g. a car can be found on a road, but not inside a playground. The model's bounding boxes for objects can compare with information in a given knowledge base deploying spatial commonsense to flag whether the output is correct or erroneous.

Likewise, the use of commonsense knowledge can prevent various object detection systems from making illogical calculations and drawing incorrect inferences by improving their training. CSK can therefore play quite a significant role in the realm of object recognition.

This can be advantageous to autonomous vehicles as well as human-robot collaboration. Various such knowledge bases, as discussed in this article, have been elaborated with respect to the extraction and compilation of commonsense knowledge in a recent tutorial [37] at the ACM conference on Web Search and Data Mining.

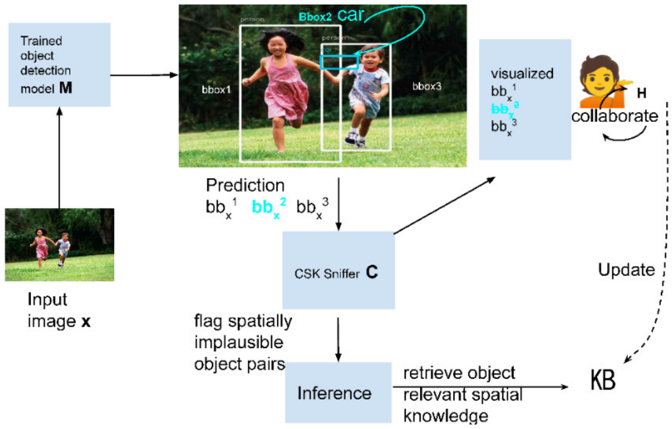

Fig. 3. CSK-SNIFFER: Web-based system to flag object detection errors via spatial commonsense [16]

3. AUTONOMOUS VEHICLES

Vehicle automation, where vehicles are fully or partially automatically controlled, is a growing research area [3, 25, 26]. The *degree of automation* entails several levels as shown in Figure 4. Level 0 implies no automation, which is primarily cruise control at most [4]. The first level, i.e. Level 1 is driver assistance, which offers dynamic cruise control and assistance with staying in lanes. However, the driver needs to control the vehicle. The second level (Level 2) is partial automation, where the vehicle can assist in controlling speed and steering. Yet at this level, the driver still needs to have some control on the vehicle. Level 3, the third level, is conditional automation where the vehicle can drive itself under ideal conditions, but requires the driver to remain behind the wheel and take over if conditions cease being ideal. The fourth level, Level 4, consists of high automation, where vehicles can fully drive themselves, but only under known use cases. Lastly, the fifth level (Level 5) is full automation, where the vehicle can drive itself under any conditions. This level typically entails full autonomy, where the vehicles make their own decisions. Such fully autonomous vehicles are yet to be prominent in the real world. As we head towards greater autonomy, the role played by IoT becomes increasingly important, leading to the realm of IoV (Internet of Vehicles). Quoting researchers in this area, “IoV evolved from Vehicular Ad Hoc Networks (”VANET”, a category of mobile ad hoc network used for communication between vehicles and roadside systems) and is expected to ultimately evolve into an Internet of autonomous vehicles” [17]. Accordingly, IoV is defined as: “a network of vehicles equipped with sensors, software, and the technologies that mediate between these with the aim of connecting and exchanging data over the Internet according to agreed standards”. Needless to say, IoV has a special place in vehicular autonomy, its demand being likely to grow in the future.

While autonomous vehicles are useful, they cause people to lose the feeling of driving. Therefore, adding *hand gesture reading* to allow for interaction between a driver and autonomous vehicles helps to improve driver experience [26]. Hand gestures are often preferable over voice interfaces since they work effectively in loud environments, in other words they are not subject to overwhelming by background noise. Gestures can include turning, lane changing, increasing the speed, reducing the speed, orienting the car and canceling inputs. The use of simple gestures is likely to reduce the task workload while improving the driving experience. The drivers have a better experience by interacting with the vehicle and feeling that they have some control. In general, this demonstrates how numerous factors are important in vehicle automation, especially for fully autonomous or semi-autonomous vehicles.

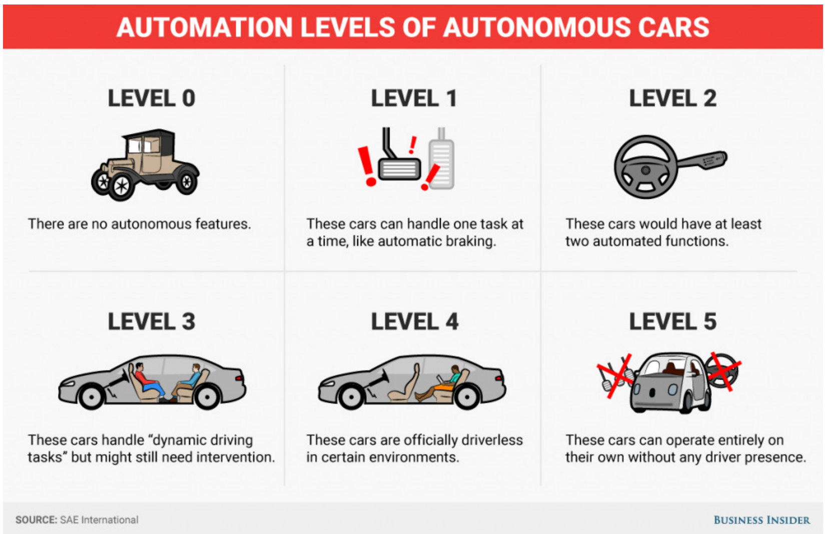

Fig. 4. Degrees of automation in vehicles, emphasized further by IoV [4]

Autonomous vehicles are not limited to cars; they can be trains and airplanes as well. There are countries where fully autonomous subway trains are running with passengers, e.g. in Singapore. In some places, there exist such trains but human drivers are still in place as a backup for safety reasons. Among planes, there are *unmanned air vehicles (UAV)* that can handle a variety of problems due to their small size and low cost [3]. Brigham Young University has developed small UAVs that can utilize an autopilot to fly to a destination. Simultaneously, a voice interface is usable to provide instructions to the UAV. Their UAVs calculate a path containing positions along with a trajectory of positions and corresponding times. Such UAVs are useful since they can fly without being guided continuously and do not need to hold a person, unlike manned aircrafts. They also do not need explicit controls by a human. These features offer the UAVs tremendous potential, making them very useful. The concept of IoV plays a subtle role here.

In order for full automation to occur, vehicles need substantial amounts of training to have their AI systems perceive the surroundings, process the perceived information, and act appropriately based on the processing. Commonsense knowledge from a myriad of Web-based repositories (such as those discussed in the previous section) can be useful to handle such challenges. Self-driving vehicles need to detect objects precisely; and CSK can refine that object detection, e.g. focusing on determining how objects are connected to the context in which they occur [29]. For example, a vehicle with one tire lower than the others may have a flat tire and can be potentially dangerous due to possibly erratic behavior. Vehicles must analyze their environment with commonsense knowledge to handle a multitude of conditions they may face.

Commonsense knowledge in conjunction with state-of-the-art Web-based object detection systems such as *YOLO (You Only Look Once)* [38] can be used to differentiate between objects based on their properties as shown in Figure 5, e.g. stationary objects and moving objects. This CSK-based detection is significant in numerous real-world applications for automated driving scenarios. For example, in 2016, a Tesla vehicle collided with a truck in motion after mistaking it for an overpass, thus causing a fatal accident [33]. A CSK-based system, analogous to a good human driver, would know that overpasses are stationary while trucks are movable. Since such a system would know that overpasses do not move, it would correctly be able to distinguish a moving truck from a stationary overpass and avoid such accidents [33]. Not only is this CSK-based information useful in the crucial moments of decision-making in autonomous vehicles, it is also stored for future reference to guide other relevant decisions. CSK is advantageous for handling adverse conditions, such as rain indicating that the road is slippery since the road is likely to remain slippery even after rain stops. Additionally, objects need to have their information stored even if they are not visible since they are still present in the detection system's environment [29]. Vehicles can detect objects more easily and undergo automation more effectively through the usage of commonsense knowledge. This is particularly useful in enhancing autonomous vehicles in order to guide their decision-making with respect to scenarios they encounter for the first time. Since they may not essentially have prior training with identical scenarios, CSK embodied within them would enable them to make better, more informed decisions closer to the thresholds of human cognition. This is in line with the challenging AI vision of autonomous vehicles being commonplace on the road in the near future.

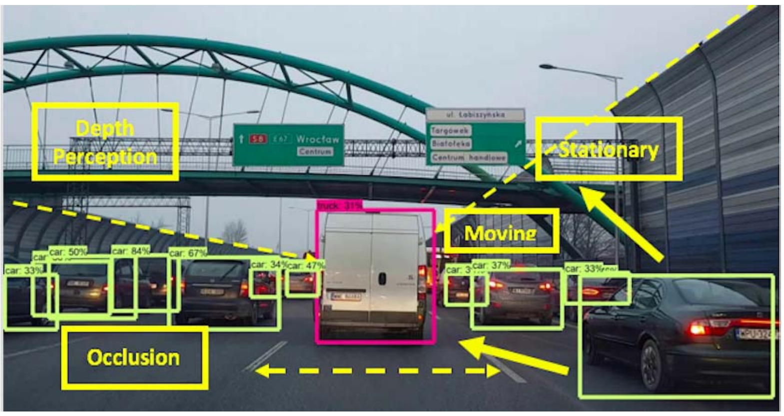

Fig. 5. CSK-based object detection with knowledge extracted from repositories on Web showing properties of objects with their bounding boxes [29].

4. HUMAN-ROBOT COLLABORATION

Human-robot collaboration (HRC) is quite important and truly beneficial in AI [10, 21, 22, 28]. Traditional industrial robots require extra guarding and equipment for safety, thereby increasing their cost and bulkiness while decreasing their flexibility. Additionally, traditional robots are not easily modifiable while the flexibility of collaborative robots can shift more seamlessly to different tasks making them more suitable in modern-day industrial settings, e.g. in the context of smart manufacturing. *Collaborative robots* can work alongside human beings within the same space [6, 8, 41]. This cooperation combines the best elements of humans and robots, such as the strength and repetition skills of robots, in conjunction with the judgement and adaptation skills of humans [41].

Human-robot collaboration is an important part of the automotive industry with significant future impacts, especially in the context of smart manufacturing. The concept of *smart manufacturing* as the very name implies entails executing the manufacturing processes in a smart an intelligent manner utilizing the latest advances in modern technology. Industry 4.0 represents ideas about how robots will be applicable in the future. The concept of Industry 4.0 uses globally available information, e.g. information disseminated through Web-enabled devices and communication tools in order to shape the manufacturing processes [2]. In particular, Industry 4.0 uses the Internet of Things (IoT) along with cyber physical systems, allowing for dynamic manufacturing [2, 45]. Robots can work on their own, performing certain tasks better suited for them (than humans). They can work 24 hours a day and be controlled remotely while human workers can continue to work throughout significant parts of the day [2]. However, the purpose of robots is not to replace humans, but rather to work alongside humans [45]. Smart manufacturing therefore focuses on utilizing human ingenuity rather than minimizing the role of humans. Hence, smart manufacturing is likely to gain even more prominence in the near future, as the roles played by robots considerably increase in manufacturing, while humans continue to play critical roles therein.

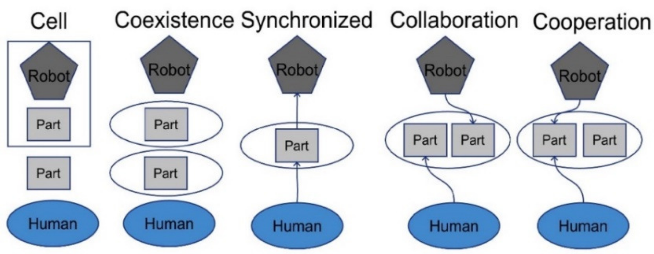

Fig. 6. Types of human-robot collaboration applicable to real-world settings and cyberspace [9]

As depicted in Figure 6, there are four main types of human-robot collaboration (HRC): coexistence, synchronized, cooperation and (full) collaboration [9, 27, 41]. In the concept of *coexistence*, humans and robots only work in the same space but without interacting with each other. For *synchronized* HRC, humans and robots do work in the same space albeit at different times, and as the name implies, they "synchronize" with each other. In the *cooperation* type of HRC, humans and robots work together in the same space simultaneously, however they work on different tasks. For the term designated as *collaboration* in HRC, typically interpreted as full collaboration, humans and robots actually work together on the same tasks, with their actions affecting each other. These different types of collaboration in the realm of HRC can be useful for different situations, since humans may perform better in some kinds of tasks while robots may work better in others. Human-robot collaboration thus has various forms and can be useful for many companies, especially in manufacturing applications. All these types of collaboration can benefit from IoT, especially in the context of Industry 4.0 and smart manufacturing. Note that such collaboration can occur with or without IoT. For example, it can occur in physical laboratory settings or in industrial applications, provided the concerned data is available in-house, along with the essential equipment. Hence, the different HRC types are pertinent to the physical world as well as cyberspace.

Robots can be beneficial for manufacturing tasks in order to increase *safety* and *productivity* in the workplace. Robots are typically used for tasks such as welding, assembly and paint spraying [23]. Robot welding is useful since it results in superior quality for products while also providing more safety and better work life for human workers. Likewise, robot assembly is useful since the costs decrease while throughput and consistency increase. Using robots can benefit humans while improving the bottom margin. Robots using paint spray not only improves consistency as well as throughput, but also keeps human operators away from hazardous environments. The use of industrial robots therefore results in more safe work environments and makes positive impacts on production as a whole. Programming robots typically involves teaching them a sequence of waypoints and events related to their actions and environments. This poses considerable issues since the programming of robots is often highly time-consuming and can require several iterations. *Augmented reality* can be used to circumvent these issues by allowing human operators to oversee a simulation of the corresponding processes and their results [34]. An augmented reality system in this context consists of a tracking system, a handheld input device, a wearable computer and the program generation. The concerned human operator would have a head-mounted display (HMD), a camera, a handheld input device and a wearable computer. The tracking here follows the HMD to accurately generate graphics and follows the handheld input device to determine its location in the augmented reality system. The wearable computer runs the algorithms that detect key fixed points; it also renders graphics and processes events, in addition to executing other functions. The human operators can begin the program generation once there are enough key points around or on the target. While doing this, the human operators see the augmented reality generated by the robotic system placed on target points. Such systems are useful since they make robot programming much faster. This is evident from the fact that in some experiments, such systems measure 5 times faster than traditional programming methods and the human operators have found the systems easier and more intuitive to use. Augmented reality systems can benefit from IoT as well.

Robots can perform better when utilizing commonsense knowledge. CSK connects spatial and temporal relationships between objects for high-level activities. Researchers at the University of Bremen, Germany have been working on a robotic system that uses CSK to execute tasks [41]. The positions of entities at a certain point of time are stored as a 4-dimensional vertex. The concerned model has a 3-dimensional image corresponding to each point in time. In addition, visual patterns detected by imaging systems connect with spatial relations. For example, 2-dimensional and 3-dimensional shapes can connect to the sizes of the respective objects and the distances between them. Having a sufficient system of spatial and temporal relationships can allow simple robot motions to handle complicated tasks. The tasks can divide into smaller parts, making it easier to handle them. Using these spatial and temporal relationships guided by CSK can aid with task execution for robots. Thus, human-robot collaboration can certainly benefit from commonsense knowledge, e.g. repositories such as WebChild [42] can be useful in this matter. While there is existing research overlapping these two areas, there is much scope for future work embodying both HRC and CSK together. This would make even better impacts in smart manufacturing contexts, since commonsense knowledge would increase the effectiveness of HRC.

Human-robot collaboration, for real lab settings as well as for simulations in cyberspace, typically requires *communication* between humans and robots. The robots need to communicate their planned actions to the humans while reading signals from humans and coordinating their actions with humans as well. Humans utilize different mechanisms in order to execute joint actions. These include cues that help with improving coordination. One such cue is *joint attention*, which involves having two entities look at the same area. Researchers have tested how different types of gaze cues from robots would affect joint attention [32]. In a given study, a robot asked human participants to move objects with different colors to boxes of different colors. Cues were either *congruent gaze cues* with the robot referencing an object and looking at it; temporally incongruent with the robot first looking at an object before actually referencing that object; spatially incongruent with the robot referencing an object and then looking at a different object; or were not provided at all. Testing ultimately found that the *congruent gaze cues* were the most effective due to the manner in which they combined gazing with vocally referencing the objects. This is evident from Figure 7. These *congruent gaze cues* resulted in the greatest competence and lowest time to locate the correct objects.

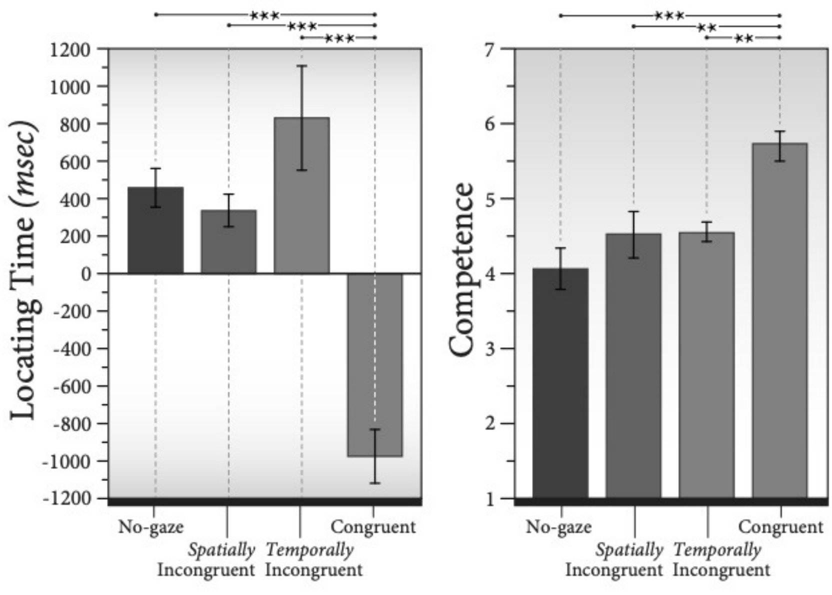

Fig. 7. Testing results for different types of robot cues in HRC [32]

Robots can additionally provide aid for tasks by repairing issues when they occur, such as a person requesting more information, a person being uncertain and hesitant, or a person making a mistake. This aid can either be provided online via suggestions and instructions or can actually be executed in a physical setting. An interesting study tested how *different forms of repairs* would correlate with the total number of breakdowns in communication [32]. Repairs are either non-existent, where the robot would not provide repair and would only add more directions after a task is completed; simple repair where the robot would responds to yes or no questions, and for other questions would repeat the instructions; or textithumanlike repair where the robot would actually provide the appropriate repair.

This study found that there were significantly fewer breakdowns when *humanlike repair* was used than with non-existent repair or simple repair [32] as shown in Figure 8. Participants additionally had the best experience with *humanlike repair*. In general, communication between humans and robots is important for designing robots in human-robot collaboration since it helps to maximize the ability of human workers. Whether the communication occurs online or in real-world settings, it has a significant impact on the effectiveness of the concerned HRC tasks.

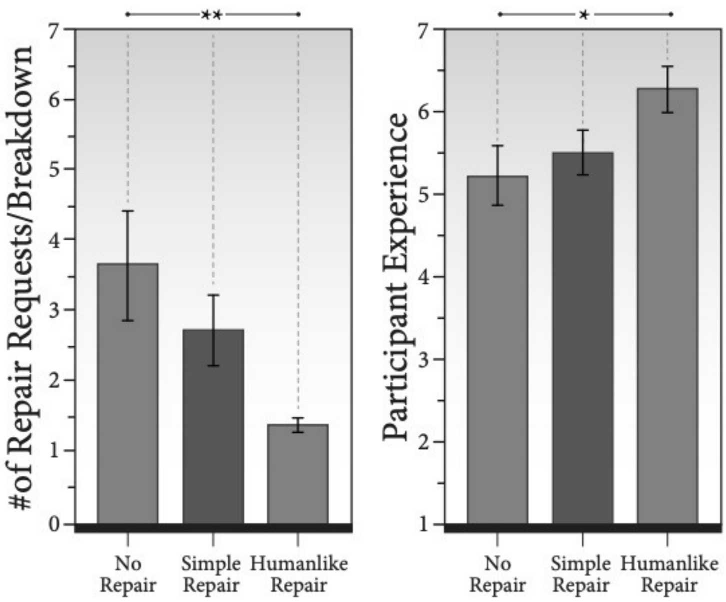

Fig. 8. Testing results for different types of repair mechanisms in HRC [32]

Human-robot interaction can improve by analyzing video games in order to form a model. These games can possibly be on standalone devices, however most of them today are available on Web-enabled devices such as smartphones. It has been observed that robots are currently prevalent with non-specialist users. Successful video games, especially those with Web access (due to its ubiquitous availability), can provide the players with the necessary information. These games are usually simple and enjoyable to control. If they are available on devices such as smartphones, their usefulness is further enhanced due to wider access, particularly the advantage of being available anywhere and anytime. In an interesting study, the benefits of a video game based interface have been demonstrated by comparing a virtual cockpit screen to an augmented reality interface [39]. The virtual cockpit screen primarily shows data points and a top-down view of navigation waypoint. On the other hand, the *augmented reality* interface shows an outline of the aircraft and a live video from the outside of the airplane. The issue with the virtual cockpit is that there are many areas to focus on while the augmented reality interface presents only the most important information and user only needs to focus on two areas. Hence, the augmented reality interface proves better as inferred from the results of this study. Since video games provide an interface with which the players interact, they can provide inspiration for designing good interfaces between humans and robots in HRC. Such interfaces, often leveraging the latest advances in Web technology and getting updated accordingly, would enhance human-robot interaction, thereby aiding HRC as a whole.

Additionally, human-robot collaboration requires effective *motion planning* and scanning [19, 31, 48]. This applies to online simulation tasks via the Web as well as to their large scale execution, e.g. in the industry. The former can be useful for educational and training purposes while the latter can be more applicable to the outputs in the given domain, e.g. the products being manufactured. Systems entailing these aspects would be useful since personalization for manufacturing helps with customization while also adhering to the goals of high productivity and low costs. While complex algorithms can be developed for this purpose, classical search methods, e.g. the A* search with heuristics, can be investigated as well [31] with reference to motion planning in mobile robotics. AI-based planning systems for robot motion planning, task planning and scheduling would help with achieving various goals. Robots need to have certain principles while collaborating. For instance, the robots should manipulate heavier objects to make the work easier for humans, they should select objects that can easily be lifted and should maintain a proper sequence for the current task [9]. Prioritization can be relevant to the actions of the humans or to minimizing the costs in general; these aspects can involve different parameters. Robots should carefully choose the correct tasks to execute in order to benefit humans.

Human-robot collaboration does not always involve one robot; it can involve *multiple robots* and thus motion planning needs to consider the actions of other robots. Motion planning involves path planning and velocity planning. A group of robots can have their motions controlled by specific algorithms as found in the literature [19]. First, each robot plans its own path and then the algorithm provides a path for each robot based on collisions. While traveling along their paths, the robots broadcast their paths in order to inform other robots. In order to avoid collisions, robots insert delays that can involve a buffer for calculating distance and decelerating in time. It is evident while testing in simulations, that the robots need to avoid colliding with each other and that this is successfully achievable. The given algorithm for controlling the robotic motion applied herewith can be useful to allow for human-robot collaboration on a larger scale. If multiple robots are in the same physical space, the presence of the Web is not mandatory. However, in today's world, multiple robots can interact across different locations due to which Web access and IoT can offer added benefits. Robots working together in cyberspace are not limited to chatbots alone. In situations where multiple robots collaborate, e.g. in smart manufacturing contexts, information can be transmitted across the Web, and the use of IoT can play an important role here. Much of this has futuristic angles.

Although commonsense knowledge based interaction and effective communication happen to be occurring, humans ultimately need to trust robots, i.e. have faith in them for collaborating with them, which makes the concept of *trust* very important to consider [5]

If human users overly trust a robot, there may be issues where the latter has too many tasks to handle. If users insufficiently trust a robot, its productivity and usefulness can potentially reduce. Human trust in robots is therefore a feature that needs adequate management such that the robot is most effective. The robot system can either ignore trust on focus on trust. The graphical plots in Figure 9 display the *estimated trust* on a 7-point scale at certain time-periods [5] For this simulated run, the robot system planning for trust waits to pick up the object (in this case, wine) until the average trust is sufficient at T=5. On the other hand, the robot system ignoring trust immediately picks up the wine, likely resulting in a substantial amount of human intervention.

For instance, in order to build trust, robots can start with low risk tasks so that they are trusted for higher risk tasks. In a study within this area, researchers had a person and a robot cooperate to clean off a table, with bottles, cans and wine glasses placed on it [5]. Robots that focused primarily on maximizing the reward removed the wine glass first, causing more human-intervention. Robots that focused on trust removed bottles before attempting to remove the wine glass, which caused the humans to be less likely to intervene. Detecting when a human overly trusts a robot is important as well, and having the robot intentionally mismanage a task is method of re-engaging a human collaborator. By managing trust correctly, robots can better help humans with various tasks.

While these experiments [5] have been conducted in a physical setting with humans and robots working in the same room, the results can be applicable to HRC in cyberspace as well. An added note here is that of the current scenario since 2020 when many operations have moved online. Some tasks that were otherwise executed in a physical setting, have transferred to the online mode with simulations as opposed to real experiments. The concept of "trust" in the real-world therefore can be mapped to cyberspace. Therefore, we claim that the findings of this study and other such studies on trust are very significant in the real world as well as the Internet world and can be useful in HRC.

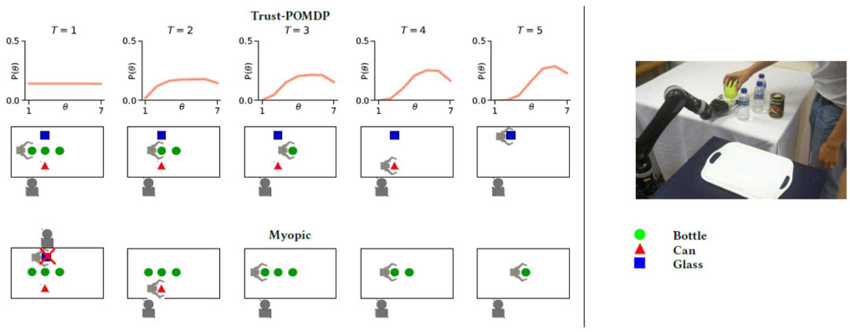

Fig. 9. A robot removing objects from a table with a human collaborating: levels of trust as depicted in the physical world, applicable to cyberspace as well [5]

Human-robot collaboration can thrive based on the utilization of commonsense knowledge and reasoning. An interesting approach in this context harnesses commonsense premises with the goal of task optimization for robot planning [8], considering the domain of smart manufacturing. It encompasses some fundamental commonsense facts along with mathematical modeling, and programs them into a system for HRC in smart manufacturing.

Commonsense premises modeled into mathematical equations in this work include the following. Firstly, humans would be inclined towards carrying lighter parts as well as and parts closer to them since that provides ease and comfort. Secondly, humans would tend to carry heavier parts more slowly than lighter parts. Thirdly, humans would prefer to handle less stable parts thereby preventing damage of parts. Fourthly, humans should not handle dangerous parts, as human safety is essential [10].

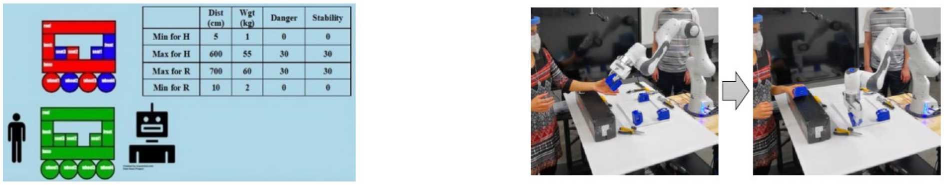

Fig. 10. Snapshots from experiments with commonsense knowledge based human-robot collaboration for smart manufacturing. Left: online simulation using vehicle assembly. Right: real lab experiments with sample vehicle parts for assembly, conducted with actual humans and robots [8, 10].

Experiments conducted using this approach [8, 10] with reference to vehicle assembly corroborate its effectiveness. These experiments include simulations as well as real-world experiments to demonstrate the deployment of the commonsense reasoning to enhance task planning in human-robot collaboration for smart manufacturing. The authors of these papers indicate that they had to resort to online simulations in their early research due to the onset of the COVID pandemic. As noticed in the literature, much work in various domains has been impacted by COVID, due to which people have shifted to online platforms, often expressing their opinions on social media [35]. Numerous opinions have been analyzed in interesting studies, in order to fathom user preferences and gain more insights into the manner in which COVID has altered trends in various aspects of life, including academia. Likewise, in this research conducted at a university [10], authors have considered online simulations initially. Further studies in real laboratory settings have been conducted at a later stage. It is important to note that the results achieved with online simulations in this work have been successfully corroborated by corresponding laboratory experiments, thereby confirming the fact that physical as well as online settings can yield similar results in HRC. Figure 10 provides snapshots from these experiments of the commonsense knowledge based human-robot collaboration approach. Experiments reveal that this approach, leveraging reasoning based on commonsense knowledge indeed achieves good task planning in human-robot collaboration, in terms of the efficiency of execution while also enhancing human safety and comfort.

Robots use policies to determine a given task and the corresponding actions thereof, based on the current world state. Robots traditionally undergo programming to handle tasks based on domain models and mathematical policies, but these approaches require defining the domain accurately. The active involvement of domain experts can be imperative to develop domain models. Instead, robots can use *learning from demonstration* (LfD), where a task is executed by a human or another robot [1]. LfD is particularly beneficial since it does not require expert knowledge and ordinary people can demonstrate how to execute the tasks. This makes LfD more flexible and intuitive than traditional robot teaching systems. There are several different LfD forms. The first is teleoperation, where the teacher operates the robot learner and the robot's sensors save the inputs. The second is shadowing, where the robot records the task execution while attempting to mimic the teacher executing the task. The third is imitation, where the robot watches a non-identical entity perform a task and this imitation is recordable through sensors on a teacher or through observation tools on the robot. The different forms of learning from demonstration can be advantageous to help robots learn how to execute the respective tasks. LfD paradigms can be implemented in cyberspace as well. Some researchers [11] put forth a study of using a robotic LfD system that is able to garner huge amounts of human-robot interaction data via a Web-based interface. They assess the impact of various perceptual mappings between the human teacher and robot on the LfD tasks. They demonstrate software tools to foster Internet-mediated human-robot interaction and collection of huge data via crowdsourcing. Their study reveals that humans are better at teaching a robot to navigate a maze when offered information limited to the robot's perception of the world.

In another study [31] that entails maze navigation, researchers investigate fundamental search methods in AI, i.e. breadth-first search, depth-first search and A* search, with respect to mobile robots. They list areas where this work is useful with respect to human-robot collaboration in addition to outlining other robotic applications. For instance, domestic robotic settings are addressed in this work, and autonomous vehicles are mentioned as well. They discuss the pros and cons of each search method as per the results obtained in their study; while A* is generally the preferred method, the others are also found to produce interesting results without the need for heuristics. This study is executed almost entirely in an online setting, with its observations being applicable to cyberspace as well as in real-world scenarios.

An interesting field where human-robot collaboration is useful is agriculture. Estimates indicate that global population will reach 9.1 billion by 2050 and that urban population will be 70% of the global population. Due to this, global food production will need to increase by 70% [47]. In this context, agricultural robots can aid in raising the food production.

Fig. 11. Workers using a robot vehicle for tree fruit farming (right, shown along with its operator panel for controls (left): such information can be disseminated via the Internet for further collaboration [47]

An example of this form of aid is a robot vehicle designed to help with tree fruit farming [47] as illustrated in Figure 11. The vehicle has three main modes, namely, Mule Mode, Pace Mode and Scaffold Mode. The Mule Mode has the robot assisting with tasks such as harvesting while following a group of workers. The Pace Mode has the robot perform a specified task over a certain area. Lastly, the Scaffold Mode has the robot traveling while acting as a scaffold on which humans can stand. The robot's usefulness is evident from the fact that workers are able to trim the trees more than twice as fast while using the Scaffold Mode versus using ladders. Thus, humans and robots working together can help in the agricultural domain. Moreover, useful information in this context can be captured and shared in cyberspace. It can be used by other farmers as well as scientists, and can be deployed to learn lessons from current contexts to enhance future decision-making. Web applications can assist with querying and drawing inferences in this respect as surveyed in the literature [13]. All of this can have the positive impact of increasing food production, thereby increasing the number of people with access to food while also decreasing any adverse environmental effects. Hence, in addition to the various contexts discussed here in human-robot collaboration that have direct or indirect impacts on smart manufacturing, there are also good impacts of HRC on related areas such as smart agriculture. This in turn, can have impacts on aspects related to urban areas, e.g. the relative growth of agriculture versus the unprecedented expansion due to urban sprawl, wherein GIS systems based on the Web can be used to access current data and make predictions for the future [44], e.g. spatial decision support systems (SDSS) can be developed and enhanced in this context.

While HRC has been studied and implemented in many real-world scenarios, many facets of HRC offer scope for future work, keeping in mind commonsense knowledge, along with Web perspectives such as IoT [18] and IoV [17]. Other Web-related aspects such as Cloud and Edge Computing [21, 22] are certainly worthy of further investigation as per their impacts on the paradigms surveyed in this article. This brings us to our conclusions along with open issues for further study.

5. CONCLUSIONS AND OPEN ISSUES

Commonsense knowledge is intuitive among humans who use it regularly, and can likewise help machines perform better if harnessed well. Vehicle automation and human-robot collaboration are areas that are quite important currently and have high potential in the future. Autonomous vehicles need to be able to interpret what they see via good object recognition techniques and must be able to handle changes as well as unseen situations through well-informed decision-making. The Web and its related facets such as IoT and IoV play crucial roles here. Autonomous vehicles thus need to avail of the advances in AI and IoV, especially facets entailing commonsense knowledge and object recognition, in order to come significantly closer to the thresholds of human cognition and enhance their decision-making capabilities. Human-robot collaboration can be beneficial to various fields, e.g. smart manufacturing entailing IoT. Robots and humans can benefit from working together, and this is observable via interesting results in various studies with demonstrations involving real-world applications.

Smart manufacturing leveraging human-robot collaboration yields good results from the productivity standpoint as well as other aspects such as human safety in the workplace. Smart manufacturing makes broader impacts on the paradigm of smart cities that is receiving much attention today. Utilizing commonsense knowledge in the areas of human-robot collaboration and autonomous vehicles proves useful in existing applications. Researchers anticipate that the further use of commonsense knowledge and reasoning will result in substantial improvements, enhancing perspectives such as smart manufacturing thereby making better impacts on smart cities.

In sum, this article provides an insight into recently growing fields relevant to robotics applications, namely, commonsense knowledge, autonomous vehicles and human-robot collaboration, especially considering Web perspectives. To the best of our knowledge, this paper is among the first works to survey all these areas together, especially considering joint work among them and discussing their impacts on smart manufacturing. This constitutes the novelty of the paper. Based on our literature survey in this paper, as well as our own research and development initiatives, we encounter a few open issues for further research in the areas discussed here. We list these as follows.

- Studying commonsense knowledge with deep learning techniques such as CNN (convolutional neural networks) and RNN (recurrent neural networks) to possibly whiten the black box as per **explainable AI for heterogeneous data** including audiovisuals, hypermedia etc.
- Exploring the specific use of human-robot collaboration in **IoT based smart manufacturing** contexts for vehicles via knowledge bases entailing commonsense concepts and reasoning
- Embedding autonomous vehicles with better object recognition using **state-of-the-art IoV and AI** such that their object detection ability is at par with or better than that of good human drivers
- Extracting **commonsense knowledge from Web-based repositories** for autonomous vehicles to enhance decision-making, especially for first time situations, with safety being the top priority
- Delving deep into the role played by **Cloud and Edge Computing** in conjunction with the paradigms of autonomous vehicles, commonsense knowledge and human-robot collaboration
- Investigating human factors such as **human comfort and safety** in human-robot collaborative tasks by building computational models, along with Web access, to optimize and improve the task quality in human-robot teams

Further research along these directions would yield even better outcomes and would thus be beneficial to applications in the areas of commonsense knowledge, human-robot collaboration and autonomous vehicles, particularly for joint work involving these paradigms, and harnessing the capabilities of the Web along with paradigms such as IoT and IoV, catering to a multitude of data types such as audiovisuals and hypermedia. This is likely to have impacts on various areas such as smart manufacturing and smart devices, hence extending broader impacts to smart cities entailing advanced robotic systems in various applications. This article can be useful to readers in several facets of AI, robotics, Web-based systems and data science. In addition to the main topics addressed here entailing commonsense knowledge, autonomous vehicles and human-robot collaboration, it would also propel other research in other related areas.

ACKNOWLEDGMENTS

The authors acknowledge the NSF grant 2117308 MRI: Acquisition of a Multimodal Collaborative Robot System (MCROS) to Support Cross-Disciplinary Human-Centered Research & Education. Christopher J. Conti acknowledges a Teaching Assistantship from the Department of Computer Science. Dr. Aparna Varde is a visiting researcher at Max Planck Institute for Informatics, Germany. She incurs support from a Faculty Scholarship Program from Computer Science and a Doctoral Faculty Program from the Graduate School. Dr. Weitian Wang has received startup funds in the Computer Science Department through the College of Science and Mathematics.

REFERENCES

[1] ARGALL, B. D., CHERNOVA, S., VELOSO, M., AND BROWNING, B. 2009. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*. doi.org/10.1016/j.robot.2008.10.024.

[2] BAHRIN, M. A. K., OTHMAN, M. F., AZLI, N. H. N., AND TALIB, M. F. 2016. Industry 4.0: A review on industrial automation and robotic. *Jurnal Teknologi 78,* 6–13, 137–143.

[3] BEARD, R., KINGSTON, D., QUIGLEY, M., SNYDER, D., CHRISTIANSEN, R., JOHNSON, W., MCLAIN, T., AND GOODRICH, M. A. 2005. Autonomous vehicle technologies for small fixed-wing UAVs. *Journal of Aerospace Computing, Information and Communication 2,* 1, 92–108.

[4] CHATURVEDI, A. 2018. "Did you know what are the five levels of autonomous cars?". *Geospatial World, Advancing Knowledge for Sustainability*.

[5] CHEN, M., NIKOLAIDIS, S., SOH, H., HSU, D., AND SRINIVASA, S. 2018. Planning with Trust for Human-Robot Collaboration. *ACM/IEEE Intl. Conf. on Human-Robot Interaction*, pp. 307–315.

[6] CHEN, Y., WANG, W., ABDOLLAHI, Z., WANG, Z., SCHULTE, J., KROVI, V., AND JIA, Y. 2018. A Robotic Lift Assister: A Smart Companion for Heavy Payload Transport and Manipulation in Automotive Assembly. *IEEE Robotics & Automation Magazine 25,* 2 (2018), 107–119.

[7] CLARK, P., COWHEY, I., ETZIONI, O., KHOT, T., SABHARWAL, A., SCHOENICK, C., AND TAFJORD, O. 2018. Think you have Solved Question Answering? Try ARC, the AI2 Reasoning Challenge. *CoRR arXiv preprint, 2018, arXiv:1803.05457*.

[8] CONTI, C. J., VARDE, A., AND WANG, W. 2020a. Task quality optimization in collaborative robotics. *IEEE International Conference on Big Data (2020)*, 5652–5654.

[9] CONTI, C. J., VARDE, A. S., AND WANG, W. 2020b. Robot action planning by commonsense knowledge in human-robot collaborative tasks. *IEMTRONICS - International IOT, Electronics and Mechatronics Conference, Proceedings (Sep. 2020).* doi.org/10.1109/IEMTRONICS51293.2020.9216410.

[10] CONTI, C. J., VARDE, A. S., AND WANG, W. 2022. Human-Robot Collaboration With Commonsense Reasoning in Smart Manufacturing Contexts. *IEEE Transactions on Automation Science and Engineering 19,* 3 (2022), 1784–1797. DOI:https://doi.org/10.1109/TASE.2022.3159595.

[11] CRICK, C., OSENTOSKI, S., JENKINS, O. C., AND JAY, G. 2011. Human and robot perception in large-scale learning from demonstration. *HRI 2011 - Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction*, (2011), 339–346.

[12] DAVIS, E. AND MARCUS, G. 2015. Commonsense reasoning and commonsense knowledge in artificial intelligence. *Communications of the ACM 58,* 9, 92–103.

[13] DRURY, B., FERNANDES, R., MOURA, M. F., AND DE ANDRADE LOPES, A. 2019. A survey of semantic web technology for agriculture. *Information Processing in Agriculture 6,* 4 (Dec. 2019), 487–501. DOI:https://doi.org/10.1016/J.INPA.2019.02.001.

[14] ELAZAR, Y., MAHABAL, A., RAMACHANDRAN, D., BEDRAX-WEISS, T., AND ROTH, D. 2019. How Large Are Lions? Inducing Distributions over Quantitative Attributes. *CoRR arXiv preprint*. 2019, arXiv:1906.01327.

[15] GARG, A., TANDON, N., AND VARDE, A. S. 2020. I am guessing you can ' t recognize this: Generating adversarial images for object detection using spatial commonsense. *(Student Abstract) . Association for Advancement of Artificial Intelligence Conference (AAAI)*, 13789–13790.

[16] GARG, A., TANDON, N., AND VARDE, A. S. 2022. CSK-SNIFFER: Commonsense Knowledge for Sniffing Object Detection Errors. *ACM EDBT (Extending Database Technology) conf. workshops*.

[17] GERLA, M., LEE, E. K., PAU, G., AND LEE, U. 2014. Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds. *2014 IEEE World Forum on Internet of Things, WF-IoT 2014*, 241–246. DOI:https://doi.org/10.1109/WF-IOT.2014.6803166.

[18] GRIECO, L. A., RIZZO, A., COLUCCI, S., SICARI, S., PIRO, G., DI PAOLA, D., AND BOGGIA, G. 2014. IoT-aided robotics applications: Technological implications, target domains and open issues. *Computer Communications 54*, 32–47.

[19] GUO, Y. AND PARKER, L. E. 2002. A distributed and optimal motion planning approach for multiple mobile robots. *IEEE ICRA (International Conference on Robotics and Automation)*, pp. 2612–2619.

[20] HAYES, B. AND SCASSELLATI, B. 2016. Autonomously constructing hierarchical task networks for planning and human-robot collaboration. *IEEE ICRA (International Conference on Robotics and Automation)*, pp. 5469–5476.

[21] HU, L., MIAO, Y., WU, G., HASSAN, M. M., AND HUMAR, I. 2019. iRobot-Factory: An intelligent robot factory based on cognitive manufacturing and edge computing. *Future Generation Computer Systems 90*, 569–577.

[22] HUANG, P., ZENG, L., CHEN, X., HUANG, L., ZHOU, Z., AND YU, S. 2022. Edge Robotics: Edge-Computing-Accelerated Multirobot Simultaneous Localization and Mapping. *IEEE Internet of Things Journal 9,* 15, 14087–14102.

[23] IQBAL, J., ISLAM, R. U., ABBAS, S. Z., KHAN, A. A., AND AJWAD, S. A. 2016. Automating industrial tasks through mechatronic systems–A review of robotics in industrial perspective. *Tehnički vjesnik 23,* 3, 917–924.

[24] KILGARRIFF, A. AND FELLBAUM, C. 2000. WordNet: An Electronic Lexical Database. *Language*. DOI:https://doi.org/10.2307/417141.

[25] LI, R., WANG, W., CHEN, Y., SRINIVASAN, S., AND KROVI, V. N. 2018. An end-to-end fully automatic bay parking approach for autonomous vehicles. *ASME DSCC (Dynamic Systems and Control Conference) Vol. 51906*, p. V002T15A004.

[26] MANAWADU, U. E., KAMEZAKI, M., ISHIKAWA, M., KAWANO, T., AND SUGANO, S. 2017. A hand gesture based driver-vehicle interface to control lateral and longitudinal motions of an autonomous vehicle. *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings (2017)*, 1785–1790.

[27] MATHESON, E., MINTO, R., ZAMPIERI, E. G. G., FACCIO, M., AND ROSATI, G. 2019. Human-robot collaboration in manufacturing applications: A review. *Robotics*. doi: 10.3390/robotics8040100.

[28] MUTLU, B., TERRELL, A., AND HUANG, C. 2013. Coordination Mechanisms in Human-Robot Collaboration. *HRI Conference, Workshop on Collaborative Manipulation*, pp. 1–6.

[29] PANDEY, A., PURI, M., AND VARDE, A. 2018. Object detection with neural models, deep learning and common sense to aid smart mobility. *IEEE ICTAI (International Conference in Tools with Artificial Intelligence)*, pp. 859–863. doi: 10.1109/ICTAI.2018.00134.

[30] PARK, T., LANK, E., POUPART, P., AND TERRY, M. 2008. "Is the sky pure today?" awkchecker: An assistive tool for detecting and correcting collocation errors. *UIST 2008 - Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, (2008), 121–130.

[31] PAULINO, L., HANNUM, C., VARDE, A. S., AND CONTI, C. J. 2022. Search Methods in Motion Planning for Mobile Robots. *IntelliSys (Intelligent Systems and Applications), Springer, Lecture Notes in Networks and Systems , vol 296*, pp. 802–822. https://doi.org/10.1007/978-3-030-82199-9_54.

[32] PELLEGRINELLI, S., ORLANDINI, A., PEDROCCHI, N., UMBRICO, A., AND TOLIO, T. 2017. Motion planning and scheduling for human and industrial-robot collaboration. *CIRP Annals - Manufacturing Technology*. https://doi.org/10.1016/j.cirp.2017.04.095.

[33] PERSAUD, P., VARDE, A. S., AND ROBILA, S. 2018. Enhancing autonomous vehicles with commonsense: Smart mobility in smart cities. *IEEE ICTAI (International Conference on Tools with Artificial Intelligence)*, pp. 1008–1012. doi: 10.1109/ICTAI.2017.00155.

[34] PETTERSEN, T., PRETLOVE, J., SKOURUP, C., ENGEDAL, T., AND LOKSTAND, T. 2003. Augmented reality for programming industrial robots. *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 319–320.

[35] PURI, M., DAU, Z., AND VARDE, A. S. 2021. COVID and Social Media: Analysis of COVID-19 and Social Media Trends for Smart Living and Healthcare. *ACM SIGWEB (Autumn 2021),* Article 5, pp. 1–20. DOI:https://doi.org/10.1145/3494825.3494830.

[36] PURI, M., VARDE, A. S., AND DE MELO, G. Commonsense based text mining on urban policy. *Springer LREV journal (Language Resources & Evaluation)*. https://doi.org/10.1007/s10579-022-09584-6.

[37] RAZNIEWSKI, S., TANDON, N., AND VARDE, A. 2021. Information to Wisdom: Commonsense Knowledge Extraction and Compilation. *ACM WSDM (Web Search and Data Mining) conference*, pp. 1143–1146. https://dl.acm.org/doi/10.1145/3437963.3441664.

[38] REDMON, J., DIVVALA, S., GIRSCHICK, R., AND FARHADI, A. 2016. You only look once: Unified, real-time object detection. *IEEE CVPR (Computer Vision and Pattern Recognition)*, pp. 779–788.

[39] RICHER, J. AND DRURY, J. L. 2006. A video game-based framework for analyzing human-robot interaction: characterizing interface design in real-time interactive multimedia applications. *ACM SIGCHI/SIGART conf, on human-robot interaction*, 266–273.

[40] SAKAGUCHI, K., BRAS, R. L., BHAGAVATULA, C., AND CHOI, Y. 2021. Winogrande: An adversarial winograd schema challenge at scale. *Communications of the ACM 64,* 9, 99–106.

[41] SUCHAN, J. AND BHATT, M. 2017. Commonsense Scene Semantics for Cognitive Robotics: Towards Grounding Embodied Visuo-Locomotive Interactions. *IEEE ICCV (International Conference on Computer Vision) workshops*, pp. 742–750.

[42] TANDON, N., DE MELO, G., AND WEIKUM, G. 2017. WebChild 2.0: Fine-grained commonsense knowledge distillation. *ACL (Association for Computational Linguistics), Proceedings of System Demonstrations*, pp. 115–120.

[43] TANDON, N., VARDE, A. S., AND DE MELO, G. 2018. Commonsense Knowledge in Machine Intelligence. *ACM SIGMOD Record*. DOI:https://doi.org/10.1145/3186549.3186562.

[44] THAMPI, PAMPOORE, A., VARDE, A. S., AND YU, D. 2021. Mining GIS data to predict urban sprawl. *arXiv preprint*. arXiv:2103.11338.

[45] THOBEN, K.-D., WIESNER, S., AND WUEST, T. 2017. "Industrie 4.0" and smart manufacturing-a review of research issues and application examples. *International journal of automation technology 11,* 1, 4–16.

[46] VARGHESE, A., VARDE, A. S., PENG, J., AND FITZPATRICK, E. 2015. A Framework for Collocation Error Correction in Web Pages and Text Documents. *ACM SIGKDD Explorations 2015, 17,* (1), 14–23.

[47] VASCONEZ, J. P., KANTOR, G. A., AND AUAT CHEEIN, F. A. 2019. Human–robot interaction in agriculture: A survey and current challenges. *Biosystems Engineering*. doi: 10.1016/j.biosystemseng.2018.12.005.

[48] WANG, W., CHEN, Y., AND JIA, Y. 2019. Which way is the best?: Evaluation and optimization of dual-arm robot path planning for human-robot collaborative tasks in smart manufacturing contexts. *ASME DSCC Dynamic Systems and Control Conference*. doi.org/10.1115/DSCC2019-9143.

[49] WANG, W., LI, R., CHEN, Y., DIEKEL, Z. M., AND JIA, Y. 2018. Facilitating Human–Robot Collaborative Tasks by Teaching-Learning-Collaboration From Human Demonstrations. *IEEE Transactions on Automation Science and Engineering 16,* 2, 640–653.

[50] WHITNEY, D. E., LOZINSKI, C. A., AND ROURKE, J. M. 1986. Industrial robot forward calibration method and results. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME 108,* 1, 1–8. DOI:https://doi.org/10.1115/1.3143737.

Christopher J. Conti is a Montclair State University graduate (class of 2020) with a Master of Science (MS) in Computer Science. He is currently employed at Spectrum Works, where he teaches web design and programming. His MS research focused on human–robot collaboration, commonsense knowledge, and automated vehicles.

Aparna S. Varde received her Ph.D. degree in CS from the WPI, Worcester, MA, USA. She is currently an Associate Professor of Computer Science at Montclair State University. During her sabbatical, she has been a Visiting Researcher at the Max Planck Institute for Informatics, Germany. Her work includes machine learning for robotics, drone data, and satellite imagery. She has been the research advisor / mentor / committee member for 10 Ph.D. students at Montclair, and refereed 4 external ones. Her research entails multidisciplinary AI, yielding more than 100 publications with projects spanning commonsense knowledge, smart cities, geoinformatics, text mining and robotics. She has four best paper awards at recent IEEE conferences.

Weitian Wang received the Ph.D. degree from the University of Science and Technology of China, Hefei, China, in 2016. He worked as a Post-Doctoral Fellow at Clemson University from 2016 to 2019. He is currently an Assistant Professor with the Department of Computer Science, Montclair State University, Montclair, NJ, USA. His research interests include collaborative robotics, smart systems, human-robot interaction, machine learning, and sensing technology.