

# A New Approach for Training Cobots from Small Amount of Data in Industry 5.0

Khalid Jabrane, Mohammed Bousmah  
LTI Laboratory, National School of Applied Sciences  
Chouaib Doukkali University  
El Jadida, Morocco

**Abstract**—Machine learning is a vital part of today's world. Although the current Machine Learning slogan is “big data is required for a smarter AI”. All Artificial Intelligence learning techniques require the training of algorithms with huge data. Collecting and storing this data takes time and requires increasing computer memory. In Industry 5.0, human-robot collaboration is a challenge for artificial intelligence (AI) and its subdomains. Indeed, integration of its domains is required. Many AI techniques are needed, ranging from visual processing to symbolic reasoning, task planning to mind building theory, reactive control to action recognition and learning. Otherwise, the main two obstacles to this natural workflow interaction are big data memorization and time Learning that grows exponentially with the problem complexity especially. In this article, we propose a new approach for training Cobots from Small Amount of Data in the context of industry 5.0 based on common-sense capability inspired by human learning.

**Keywords**—Small data; industry 5.0; common-sense capability; machine learning

## I. INTRODUCTION

Despite its economic progress, all industrial revolutions had an impact on the labor market with the goal of replacing human labor with machines. The first industrial revolution replaced manual work with the invention of a steam engine and the second industrial revolution enabled mass production using electric energy. The tertiary industrial revolution started the automation era with informatization based on computers and the Internet. In the fourth industrial, the super intelligence revolution based on the Internet of things, cyber-physical system, and artificial intelligence (AI) will greatly change human intellectual labor. Fig. 1 depicts the five industrial revolutions.

Industry 5.0 will transform the labor market by emphasizing the central role of humans and encouraging collaboration between humans and a new generation of robots known as "collaborative robots" or "cobots". These cobots are designed to work alongside their human counterparts and, more importantly, help them accomplish common tasks in real world. They are user-friendly and their key function is to provide physical assistance to operators by performing unpleasant and risky activities. With the introduction of Cobots there should be no fear of losing the production line due to automation which has been a major concern of Industry 4.0, as a result better agility will be added to the smart factory.

In the context of Industry 5.0, AI and cobotics must play a central role to improve the capabilities of Cobot. The cobotics is a major discipline that focuses on collaborative robots and their uses as technical agents. Moreover, Cobotics seeks to extend beyond the isolated faculties of humans and robots. Synergy is an essential factor in increasing the respective capacities of man and machine.

However, this promising vision of cobot, driven by AI and cobotics, requires significant R&D progress. Many technical challenges remain in all subfields of AI application. The neural networks of deep learning models require exposure to huge amounts of data to learn a task. Training a neural network to recognize an object, for example, could require feeding it as many as 15 million images. Acquiring relevant datasets of this size can be costly and time-consuming, which slows the pace of training, testing and refining AI systems. Furthermore, some fields suffer from a lack of data to feed a starving deep learning model.

Researchers are working hard to find ways to train systems with less data and are optimistic that they will find a viable answer. As a result, AI specialists anticipate that the "big-data" variable's tendency will be reversed in the AI growth equation. Small data will supplant Big Data as fresh and innovative AI drivers in Industry 5.0.

The goal of this study is to present a novel approach for training cobots with small amounts of data. We place a premium on a framework built on common sense and on-the-fly multitasking techniques. As a result, this article presents a summary of current cobotics research. We define cobotics and provide a brief classification of cobotic systems. The second segment discusses the main challenges of artificial intelligence, while the third section looks for potential solutions. Finally, we'll present our model.

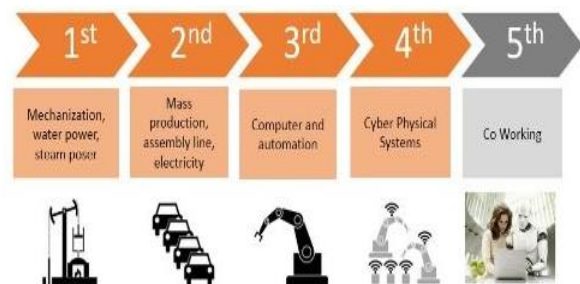


Fig. 1. The Five Industrial Revolutions.

## II. RELATED WORK

Cobotics is a transdisciplinary technology and is the intersection of three main fields such as robotics, ergonomics and cognitive. This new discipline focuses on studying the interaction of human and cobot.

Cobotics is not included in robotics because certain aspects, such as the human representation of the robot, ergonomics of the workstation, or operator acceptance, are the result of ergonomic and cognitive engineering studies.

At the end of the 1990s, the term "cobot" was describe as a passive effort-assistance device controlled directly by an operator [1]. Since then, the meaning of the term has evolved. It was only popularized in the 2010s. Current cobotics allow the extension of human gesture or behavior and is distinguished from simple robotics by the real, direct or remote interaction between a human operator and a controlled or pseudo-autonomous robotic system. This cooperative robotics is more "user-centric" [2].

While is still conceived today as a cooperative augmentative robotics, cobotics corresponds generically to the use of mechanical or artificial sensorimotor support systems developed specifically for a given task or relationship. It then becomes a form of parallel robotics or extension robotics, allowing an increase in performance, in strength, speed or precision [3].

Given the novelty of the theme, the majority of articles in the literature treat this new industrial revolution as being the era where Man and cobots will have to work side by side. Although this observation is part of our daily life in some pioneering industries in this field, this cooperation takes place for very specific and low-level tasks.

For some authors, Artificial Intelligence needs a theory based on the tasks [4], others think that to design new factories in Industry 5.0 with a human-centered perspective, technological engineering should be centered on values and ethics [5].

Other authors have even proposed probabilistic models to infer intention in human-robot interaction [6]. In this context, intention inference has been the subject of several studies and research based on the Markov model which predicts and models human behavior by a series of discrete states and actions.

Also, some studies explain an inspiring technique based on Reverse Reinforcement Learning (IRL) [7] [8]. Ordinary reinforcement learning involves the use of rewards and punishments to teach a behavior to an intelligent agent, in IRL the process is reversed, the robot observes a person's behavior to determine the goal that behavior appears to be aiming for. The problem in IRL is to determine the optimized reward function, by the agent, which best and in a transferable way defines the intended task. This approach has enabled state-of-the-art advances in several areas of robotics.

Although some articles suggest possible research avenues, finding a relevant issue linking artificial intelligence and Industry 5.0 seems philosophical if we move away from technical reflections based on the adoption of a stochastic

modeling approach, which also remain restrictive given the input hypotheses and the cases treated, which do not promise generalization.

In short, according to scientists and industry, the trend in the era of industry 5.0 is converging on the search for tools and technologies that propel the advancement of symbiotic interaction in the workplace.

In this context, recent advances in reinforcement learning have successfully combined deep learning to make significant improvements in the formation of an agent. Despite the impressive performance of Deep Reinforcement Learning (DRL) techniques on individual tasks, training a single DRL agent to perform multiple tasks remains difficult [9]. Traditional learning algorithms that consist of directly training a DRL agent for multiple tasks one by one have been shown to offer poor performance and may even fail on some tasks.

Unlike DRL which teaches an agent to perform a single task, the new multitasking DRL techniques advocate that the agent learns a single control policy that could work well on several different tasks.

The current AI approach is to train agents using as much data as possible while also blending real and synthetic data. Big data alone will not be able to overcome the difficulties of human-machine collaboration in real-world interactions.

We typically think of "big data" when we hear "artificial intelligence," because the most notable advancements in AI have been built on massive data sets. Image classification, for example, has made significant progress in the construction of ImageNet, while the GPT-3 language model has been trained on hundreds of billions of words of online text using deep learning techniques to produce human-like writing. It is therefore not surprising to see AI being closely linked to "big data" in the popular imagination. But AI is not just about big data sets, research on 'small data' approaches is being developed nowadays.

The common sense is the natural ability to make good judgment to behave in a practical and sensible way.

Common sense is the unconsciously acquired knowledge that all humans have since birth. This common-sense knowledge is gained through experience and curiosity, sometimes without the learner's knowledge.

According to John McCarthy, father of AI, Common sense knowledge includes the basic facts about events and their effects, facts about knowledge and how it is gained, facts about beliefs and desires. It includes the basic facts about material objects and their properties [10].

Common sense is the ability to interpret a situation in light of its context, based on millions of interconnected elements of common knowledge. The capacity to use this knowledge wisely qualifies the ability to perform common sense reasoning. Common sense reasoning is a central part of intelligent behavior.

In 1959, John McCarthy proposed common sense reasoning in the form of a theoretical program named Advice Taker. However, despite recent advances in machine learning,

there has been no progress in terms of true common sense reasoning skills. The recent surge in popularity of the subject can be attributed to recent advances in NLP and the importance of the task.

Since John McCarthy's hypothesis, proposed in the 1950's, that common sense programs could be developed using formal logic, Today's primary approaches to common sense reasoning in AI, as well as their taxonomies, are depicted in Fig. 2.

McCarthy's hypothetical proposition has sparked a flurry of research into logic-based approaches to commonsense reasoning. There are several research efforts; here are a few of them: situation calculus, naïve physics, default reasoning, non-monotonic logics, description logics, and qualitative reasoning, less formal knowledge-based approaches.

The 'Cyc' Project remains the most notable effort adopting the knowledge-based approach. The 'Cyc' Project spent 35 years codifying common sense into an integrated logic-based system. This awesome effort covers vast areas of common-sense knowledge and incorporates sophisticated logical reasoning techniques.

For a variety of reasons, including the fragility of symbolic logic, 'Cyc' was unable to realize its goal of delivering a generally useful common-sense service.

Artificial intelligence (AI) systems lack common sense knowledge. Furthermore, despite years of effort, developing common sense reasoning AI systems has always been a tedious task. Today's AI researchers agree that the most difficult problem in AI research is developing programs with common sense capabilities.

Formalizing common sense knowledge for any reasoning problem, no matter how simple, is a huge task. This is because common sense knowledge is implicit, whereas expert/specialist knowledge is usually explicit. As a result, developing an AI common sense reasoning system will necessitate explicitly expressing this knowledge.

An intelligent system lacking in common sense will struggle to understand its surroundings, interact naturally with people, respond appropriately in unexpected situations, or learn new experiences.

The concept of common-sense reasoning is still challenging for AI specially in the context of Human-Cobot collaboration in the industry 5.0 era. Progress in common sense applications for AI is insufficient. The difficulty lies in explicit formulations of what is common sense because it is an unstructured and very confused field.

### III. STATE OF THE ART AND TECHNICAL CHALLENGES OF AI IN THE ERA OF COBOT IN INDUSTRY 5.0

#### A. Classification of Cobot System

Cobotic systems are very diversified and their applications are numerous. Their classification is based on various and dispersed criteria. Many authors have tried to classify these cobotic systems to well structure this field.

First, the social-based classification of robot properties includes form, modality, social norms, autonomy, and interactivity [11]. Then safety-based classification of robot which was a major concern since the beginning of industrial robotics, standards and norms were established to regulate this field [12].

Moreover, Robots can be classified according to their architecture, degree of autonomy, mobility, transport capacity, handling, size, and distance from the operator. This approach disregards the human component of the cobotic system. Jean Scholtz proposes a role-based classification to more clarify the role and the nature of the Human-Robot Interaction (HRI) as described below [13]:

- Supervisor Interaction: The supervisor's role is to monitor and control the overall situation.
- Operator Interaction: The operator intervenes to change the internal model or to parameterize the software when the robot behavior is altered.
- Coworker: It works in the same environment as the robot, in parallel, and sometimes has some interactions with him, for example taking the piece he just processed.
- BYSTANDER: Present in the same environment as the robot and sometimes enters its working area, it has no real interaction with it. The robot is equipped with a presence sensor, as soon as a human enters its zone, it automatically switches to a slower mode or stops temporarily. Collaboration is reduced to a minimum.

#### B. Type of Human-cobot Industrial Collaboration

Human-robot collaboration in the industrial field can take several forms. Collaboration takes place either in a shared workspace without direct contact or in a shared workspace with direct contact. Tasks are executed with or without synchronization. The robot adapts its movement to meet the human workers, and in some cases, adjusting the movement is recommended in real time.

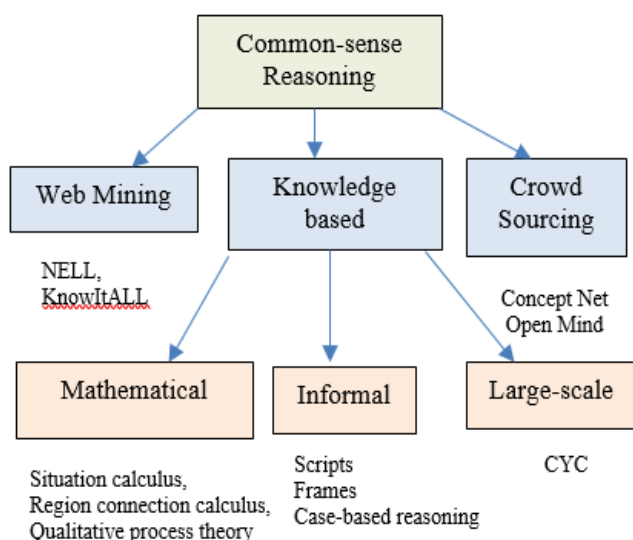


Fig. 2. Taxonomy of Approaches to Common Sense Reasoning.

According to the International Federation of Robotics (IFR), collaborative robotics today is characterized by applications where tasks run sequentially in shared workspaces, in which the robot and the employee work side by side.

Fig. 3 from IFR web site shows different types of industrial robots collaboration.

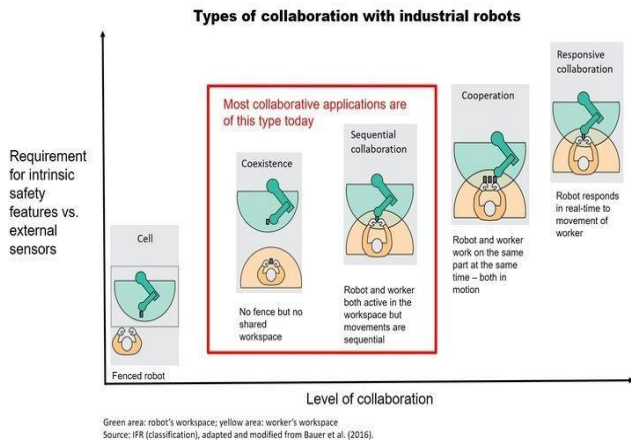


Fig. 3. Type of Collaboration with Robots Industry.

The robot frequently performs tedious or impractical tasks, such as lifting heavy parts or tightening screws. Otherwise, real-time applications in which the robot must react and expertly modify its actions to those of a worker are more technically challenging. However, robot movements are completely unpredictable; the user must ensure that the robot's potential setting meets safety requirements.

In fact, reactive human-robot collaboration will not be reachable so early in most manufacturing sectors where precision and repeatability are required to increase productivity. Otherwise, the most advanced research projects are all categorized as “sequential” or “support” collaboration scenarios [14]. Interdependent and collaborative scenarios between humans and machines require more sophisticated systems and solutions. Indeed, Cobots need a stronger semantic knowledge of the task's goal, as well as the behaviors and intentions of their human coworkers. Humans must also be able to communicate intuitively with the cobot.

### C. Technical Challenges of AI in the Era of Cobot in Industry 5.0

Today's researchers are trying to push the boundaries in order to create more advanced or complex forms of interaction by arming cobots with comprehension and anticipation skills aided by Artificial Intelligence.

Future cobots should be able to recognize human signals, movements and intentions, as well as distinguish between intentional and unintentional gestures related to the collaborative work. The natural collaboration between humans and cobots requires that cobot be able to capture, process and understand human demands with precision and robustness.

However, the technical challenges of AI lie in the interaction modalities such as speech, gaze or gesture planning

as well as motion control that must be performed in real time to ensure a natural workflow interaction. This natural workflow interaction will not be achieved only with both the classical sense-plan-act architectures and Reinforcement learning models that constitute the current state of the art in applied robotics.

Human-robot collaboration is a challenge for artificial intelligence (AI) and its subdomains. Indeed, AI's methods need to be further strengthened for a better integration of many techniques such as visual processing, symbolic reasoning, task planning, reactive control, recognition of actions and learning.

Otherwise, the main two obstacles to this natural workflow interaction are big data memorization and time learning that grows exponentially with the problem complexity especially.

1) *AI and big data:* The future of cobotics depends on artificial neural networks and deep learning which are designed to acquire advanced learning skills without the need for any type of programming. The goal of this extremely complex discipline is to enable robots to mimic the ability of humans to smoothly integrate inputs with motor responses, even as they undergo changes in their environment.

All Artificial intelligence learning techniques require the training of algorithms with huge data. Collecting and storing this data takes time and requires increasing computer memory. However, for cobotics, deep learning is a future goal rather than immediately achievable given that it requires truly massive amounts of processing power and data.

For example, deterministic problem optimization methods, including Q-learning, require recording important statistical data. Research has established that the convergence of the Q-Learning function has been proven for an infinite time. That's why it's inconceivable to teach a machine in 10 years what humanity has learned over millions of years.

One of the biggest problems of artificial intelligence is acquisition and storage. In the industry the input data comes from sensors. To validate the AI system, a mountain of sensor data must be collected. Irrelevant and noisy data sets can cause obstructions because they are difficult to store and analyze.

In addition, the AI algorithm becomes stronger and more powerful as the data collected is of good quality, relevant, and increases during its processing. The AI system fails badly when it is not fed with sufficient and good quality data; however, small variations in data quality have large consequences for results and predictions. Again, adoption of AI systems is limited for some industry sectors where data availability is insufficient.

2) *AI and multitask:* Deep reinforcement learning (DRL) has significantly improved the performance of intelligent agents. Although DRL approaches can improve agent performance to a greater extent, they were limited to systems that learned a specific task primarily through reinforcement learning algorithms.

DRL is a type of reinforcement learning and deep learning and can be defined as a crossroads between traditional and true

artificial intelligence, as illustrated in Fig. 4. The DRL combines action-based reward techniques from reinforcement learning with the concept of using a neural network to learn feature representation from deep learning.

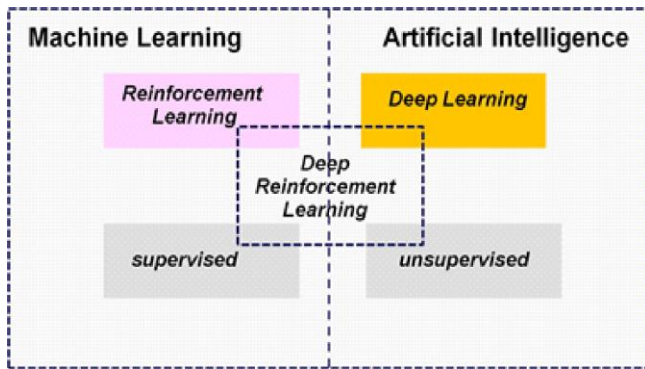


Fig. 4. Deep Reinforcement Learning.

Simultaneously, this method has proven to be inefficient in terms of data, particularly when reinforcement learning agents should interact with more sophisticated and rich data environments. This challenge emanates mostly from the limitations of deep reinforcement learning algorithm to deal with multiple scenarios of related tasks in the same environment.

Training reinforcement learning algorithms is typically time consuming, and processing requires high number of data samples to achieve an acceptable result. Second, reinforcement learning is a task-specific. Learning generalization to other tasks is practically impossible. The trained agent is deployed only on the task for which it was trained.

The section that follows provides an overview of the various challenges associated with multitasking learning in the context of deep reinforcement learning environment.

- Scalability

Scalability is a major issue in artificial intelligence when implementing multi-task learning via deep reinforcement learning [15].

One of the main shortcomings of traditional RL algorithms is their inability to extend their learning to other scenarios.

To converge to an acceptable result, RL algorithms often demand a larger number of training data samples and a longer training time [16]. There should be continuity and scalability in multitask learning by transferring the acquired knowledge to other tasks or processes. It should not take  $N$  times more samples or training time to learn  $N$  different tasks.

- The distraction problem

Balancing the demands of several tasks for the limited resources of a single learning system in a particular environment is one of the most challenging aspects of multitasking deep reinforcement learning.

Therefore, learning algorithms are often distracted to solve only a few tasks among others (a phenomenon known as the distraction dilemma) [17].

- Fractional observability

Observations made by an RL agent in many real-world scenarios are partial. Capturers only reflect a small part of the complete state of the environment [15]. When the state-action space is large, this challenge includes learning and remembering a compact representation of the environment with the most pertinent details from the environment.

- Real-world exploration

Continuous Reinforcement learning is generally based on trial and error. Exploration and exploitation are methods used by RL agents to learn by experimenting numerous possible actions from a given state in order to find the best action that delivers the best overall future reward [15]. When applying reinforcement learning to real-world problems, it is frequently difficult to achieve a higher level of exploration.

- Catastrophic interference

The goal of multitasking deep reinforcement learning (DRLM) is to train agents to learn a series of tasks with the ability to transfer knowledge from previous tasks to new tasks in order to improve the convergence speed [18]. This is a lifelong learning situation in which deep neural networks unexpectedly lose knowledge learnt in a previous task that is applicable to a new task.

The fundamental cause of this problem is changes in network parameters (weight) associated with a task that are overwritten to accomplish the goals of the following task [18]. This phenomenon is considered as a crucial obstacle to the development of a General Artificial Intelligence (AGI), as it has a negative impact on the ability to continual learning.

3) *AI and Human-Robot Interface (HRI)*: Human-Robot Interaction (HRI) is a field that studies, designs, and evaluates robotic devices for use by or with humans. The Human-robot interface (HRI) is related to the interaction modalities between the user and the robot. The sub-domains more concerned by AI research in robotics is the cognitive HRI (cHRI) that analyzes the information flow between the user and the robot and focuses mainly on multimodal interactions including textual, vocal and gestural interfaces.

Despite the advances of the AI, Human-Robot Interaction (HRI) continues to be a challenge for artificial intelligence (AI) and its subdomains. Human-robot interaction involves many technical challenges both on the technical level as well as human-centered aspects. The latter includes issues such as expectations, attitudes and perceptions.

HRI research deals with a wide range of issues, including the direct usage of robot systems that interact with humans in specific situations. The main research challenges in the field of HRI concern multimodal sensing and perception, design and human factors, and those related to developmental robotics:

- Multimodal perception

Real-time perception and management of uncertainty in detection are two of the most difficult challenges in robotics.

The need to perceive, understand, and respond to human activity in real time makes sensing and perception more complex and challenging in the field of HMI.

Human interaction sensors are far more diverse than those used in most other robotic fields today. The processing of real-time data from HRI inputs such as vision and speech poses significant challenges. Face expressions [19] and gestures [20] are examples of possible inputs that computer vision algorithms must be able to process. Correspondingly Language understanding and human-robot communication systems are still unsolved research difficulties [21, 22]. Understanding the relationship linking visual and linguistic inputs [23] then combining them to improve detection [24] and expression [25] is even more difficult. The Fig. 5 below illustrates the information processing of cobot.

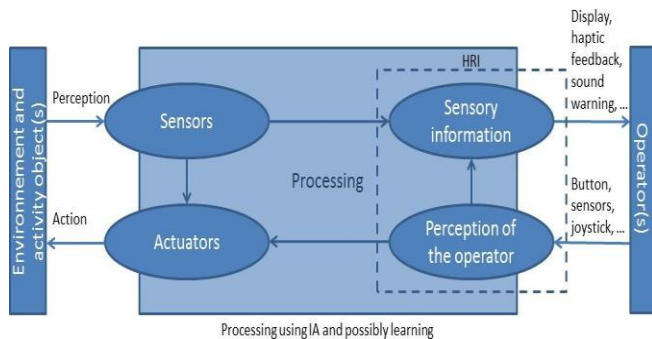


Fig. 5. Information Processing for the Cobot.

- Developmental Robotics

Developmental robotics is not a direct subset of the HRI field, but the two fields overlap significantly in their goals and have a lot in common when it comes to information acquisition techniques and multimodal perception.

The robot learns to model its environment, the objects that surround it, its own body, it learns elements of language, all this in strong interaction with the physical world but also through social interactions with humans or even other robots. The model that preoccupies the artificial intelligence researcher is no longer the chess player, but being able to learn and develop cognitively.

Developmental robotics, proposes to focus not on reproducing an immediately intelligent robot, but a robot that will be able to learn, starting with a reduced amount of innate knowledge.

#### D. Overview of Existing Solutions

DeepMind and OpenAI as research organizations have contributed significantly to the field of multi-task deep reinforcement learning (MTDRL). Their research efforts have resulted in three major MTDRL solutions, namely DISTRAL (DIStill & TRAnsfEr Learning), IMPALA (Importance Weighted Actor-Learner Architecture), and PopArt, which should be briefly mentioned in this paper as potential solutions to the issues listed in this part.

- DISTRAL (DIStill & TRAnsfEr Learning)

According to authors Nelson Vithayathil Varghese and Qusay H. Mahmoud [26]. The transfer-oriented method consists in sharing neural network parameters across related tasks in a given environment. This method has been considered as the reference for multitasking in reinforcement learning [27]. This approach encounters issues that have an impact on the learning process, such as negative knowledge transfer and ambiguity when designing a reward system for various tasks. [28]. The rewards system, in a multitasking context, should be built so that no task has to control or monopolize the shared model's learning.

DISTRAL was created as a framework for learning multiple tasks at once. DISTRAL is a novel approach to multitasking training that addresses the issues raised above.

The design's primary goal was to develop a general framework for distilling centroid policy and then transferring common behaviors into reinforcement learning across multitaskers rather than sharing parameters among the various workers in the environment [29].

Fig. 6 illustrates the Distill structure which provides a high-level view involving four tasks. The method is founded on the concept of shared policy (shown in the center), which distills common behaviors or representations from task-specific policies [30, 31]. The distilled policy is then regulated in order to direct the task-specific policies.

The knowledge gained in one task is distilled into the shared policy, which is then applied to other tasks.

The DISTRAL approach has proven to be very effective compared to the traditional method, transfer learning in multitasking therefore consists in sharing parameters over the neural networks.

DISTRAL algorithms learn faster and achieve better asymptotic performance. They are much more robust to the settings of the hyperparameters. Learning is more stable than with multi-tasking A3C baselines.

A3C was originally conceived as an extension of the actor-critic approach whose model is illustrated in Fig. 7. Two distinct neural network components are used: the actor and the critic, each having its own loss function. According to RL approaches such as Q-learning or REINFORCE, an actor can be considered as a function approximator that guides the way to act.

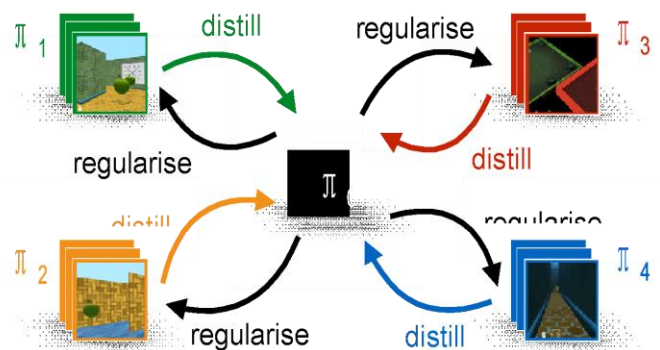


Fig. 6. Illustration of the DISTRAL Framework.

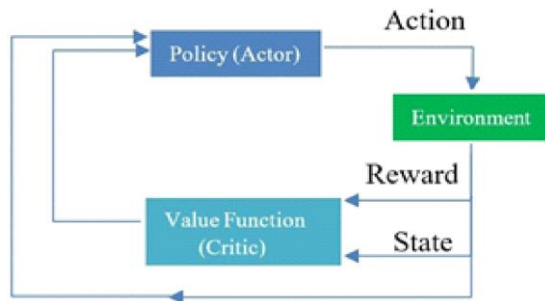


Fig. 7. Actor-Critic Model.

- Asynchronous Actor-Critic Advantage (A3C).

A3C (asynchronous actor-critical advantage) is an algorithm introduced by DeepMind. A3C offers a parallel training approach where multiple agents (called workers) run on multiple instances of the same environment [32].

Global value function is updated asynchronously by multiple workers operating in parallel environments. During the training, each parallel agents will experience a variety of different states at any time step  $t$ . The agents learning becomes nearly unique. This A3C uniqueness factor provides an effective and efficient way for agents to explore the complete state space in a given environment [33].

The role of the critic consists in evaluating the effectiveness of the policy put in place by the actor and contributes to its improvement [32].

Fig. 8 is the representation of a typical actor-critical model on which A3C is based.

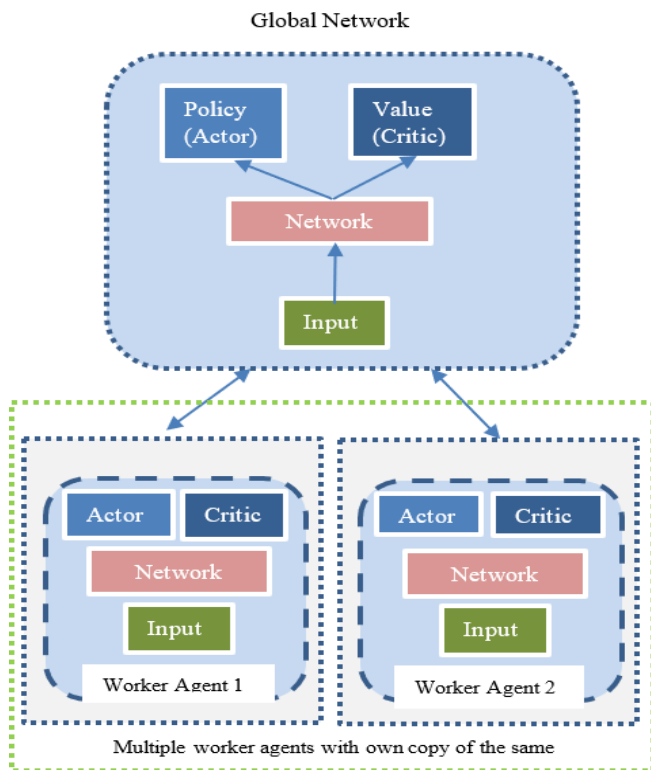


Fig. 8. Actor-Critic Model.

- IMPALA (Importance Weighted Actor-Learner Architecture).

Dealing with the increased amount of data an agent must handle, as well as the training time required, is one of the major issues in achieving functionality with a single reinforcement learning agent.

DeepMind has proposed an architecture called IMPALA to solve the abovementioned aspects of multitasking in the field of reinforcement learning. The IMPALA distributed agent architecture is based on a single reinforcement learning agent with a single set of parameters [30]. The main feature of the IMPALA approach is the ability to efficiently use resources in a single machine learning environment while scaling to many machines.

DeepMind also introduced a new correction policy based on a method known as V-trace, which enables relatively stable and fast learning by combining action and learning without compromising data efficiency or resource usage [30].

Generally, the architecture of a deep reinforcement learning model includes a single (critical) learner who is linked to several actors. Each individual actor in this ecosystem generates learning cycle parameters (also known as trajectories), which are sent as knowledge to the (critical) learner through a queue.

The learner collects the trajectories from all the other environmental actors to prepare a central policy. The policy parameters are updated with the learner (critical) and transmitted for each actor who retrieves them before the start of the new learning cycle (trajectory).

The IMPALA topology connects multiple actors and learners who should work together to build knowledge. Fig. 9 and Fig. 10 dropped from [34] depict, respectively the configurations of an IMPALA ecosystem architecture with a single learner and multiple learners.

- PopArt

Recent advancements have shown that reinforcement learning can outperform human performance in specific tasks. A specific aspect of reinforcement learning is training agents for one task at a time, learning an additional task requires instantiation of the agent [17].

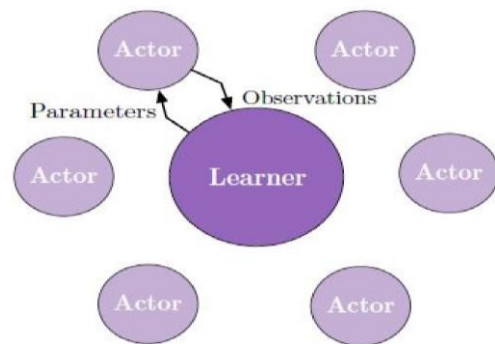


Fig. 9. IMPALA Single Learner Model.

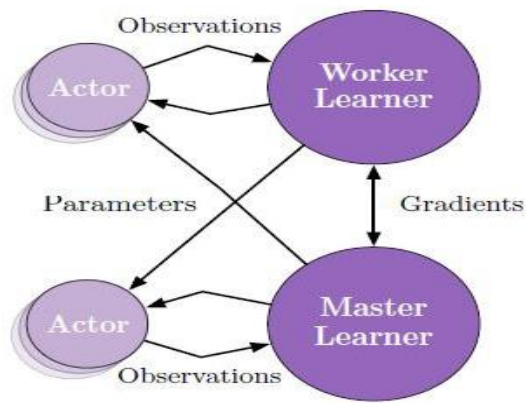


Fig. 10. IMPALA Multiple Synchronous Learners Model.

In order to overcome this limitation, much research has been carried out to improve RL algorithms by giving them the ability to carry out multiple sequential decision tasks at the same time. These research attempts that aim to support multitasking learning have often been faced with various challenges.

In general, this situation requires the establishment of a multitasking reinforcement learning (MTRL) system with strong immunity to the dilemma of distraction. Balancing the mastery of individual tasks is also important in order to achieve the ultimate goal of better generalization of learning [17]. The primary cause of the distraction scenario is that some tasks appear to be more important to the learning process due to the density or magnitude of the rewards given to them (rewards in the task).

As a result, the algorithm prioritizes these important tasks over others, sacrificing generality in the process [17].

PopArt is a new method proposed by DeepMind to improve reinforcement learning in multi-task environments. PopArt aims to reduce distractions and thus stabilize learning in order to facilitate the use of multitasking reinforcement learning (MTRL) techniques.

The PopArt method's main feature is the modification of the neural network's weights based on the results of all tasks in the environment. PopArt estimates mean and distribution of the ultimate targets for all tasks considered in the initial phase. The estimated values are then used to normalize the targets before updating the network weights. This method improves the stability and robustness of the learning process.

#### IV. PROPOSED MODEL

Artificial systems with common sense are generating a lot of interest in various fields of cognitive science and artificial intelligence to engineer common-sense reasoning into artificial agents in ways inspired by human reasoning.

Despite recent advances in many areas, artificial systems are still unable to comprehend and act on the world in a human-like manner, and are incapable of performing basic common sense thinking at the level of even young children.

So, how to concept a structure made of sense and common-sense blocks that allows cobots to understand, interact,

distinguish, and make decisions in order to overcome the challenges of the world around them. How to give cobots the intelligence and common sense they need to learn from raw and optimal representations of the scenes that fill the workplace with its objects, agents, events, and their properties.

So far, no approach has succeeded in implementing an intelligent, common-sense system. Isn't it time to reverse the trend and employ techniques and tools based on small data sets and the integration of a common-sense computational model tailored to each area of interest?

#### A. Small DATA is the Future of AI

Human intelligence has always been able to innovate and discover even before the advent of big data. All scientific discoveries throughout human history have been fueled by small amounts of data. It is estimated that 65% of these discoveries were the result of compiling a small amount of data in the form of rules, hypotheses and theories that were sophisticated and successful.

Today's biggest obstacle facing companies in developing AI systems is the lack of big data. These companies do not have the capacity of giants like Google or Facebook that rely on billions of data resources. Google, for example, benefits greatly from its massive amount of data. It can develop algorithms by processing over 130 trillion web pages, but a corporation may only have 30 relevant instances to automate an internal operation. The gap in AI adoption by companies is due in part to disparities in data resources.

Furthermore, developing a big data initiative within a company necessitates time, money, and expertise. The principles for implementing such a process include creating a data-driven program with architecture and infrastructure appropriate for the initiative's overall lifespan. The cost of this process is prohibitively expensive, and it grows in proportion to the scale of the issue and the complexity of the data.

Although it appears that current AI developments rely primarily on big data, we forget the value of observing small samples. AI becomes even smarter and more powerful if it can be trained with small amounts of data. The ultimate purpose of AI should be mastering knowledge rather than processing data. It is all about teaching a machine the knowledge it needs to complete a task.

In fact, small data mastery is essential to advance AI especially for specific industrial domains where man and robot have to collaborate by exchanging information although they are small quantities but relevant for the accomplishment of common tasks. As a result, the development of new AI techniques that do not rely on the well-known "big data" variable as input becomes critical.

The genesis of AI was to create machines capable of imitating human intelligence. However, Humans can learn from small amounts of information, they do not need to observe millions of examples of cars to learn to detect them correctly.

On the other hand, specialized learners have the ability to learn from small data because they have adequate inductive biases. Inductive biases represent knowledge of the world in



which learning will take place and are present in the model even before training begins. In other words, the model must already be capable of extracting meaning from a particular dataset. A In fact, machine learning model will learn successfully from small data only if it has a sufficient amount of this knowledge.

To summarize, small data has the advantage of being easy to collect, simple to process, ubiquitous and quickly exploitable by machine learning models. Furthermore, when applied to deep learning methods, the model can predict and converge rapidly towards an expected result even more in a narrow area, due to the ability to assign weights and rewards that will become less complex.

Investing in small data appears to provide a significant benefit because it increases the possibility of implementing alternative learning techniques.

To argue this point of view, we highlight some emerging AI tools techniques that rely only on small data and perform better than those that work with big data. These techniques appear to reinforce traditional machine learning modeling approaches and they include:

- Collaborative Machine Learning (CML).

As a distributed learning technology, collaborative machine learning (CML) trains multiple agents in a network to build a common and robust machine learning model without sharing data.

Similarly, to the synergy of heterogeneous human teams, task offloading allows agents to hold different, complementary and private representations of the training environment. The peripheral agents' joint learning is then achieved by parallelization and co-inference of distributed model learning.

The goal of collective machine learning is to create a unique predictive model that is more accurate than the sum of its parts.

- Few-Shot Learning

Few-Shot Learning is a technique that consists of performing supervised classification or regression based on a very small number of samples. Few-Shot Learning (FSL), also known as low intensity learning or spot learning, is a machine learning model that serves primarily as a tool for training machine learning algorithms with data relevant to the training context, even if it is small in number.

Else, Few-Shot Learning is different from standard supervised learning, which trains a model to recognize images in the training set and then generalize to the test set. In contrast, the goal of this technique is to distinguish similarities and differences between objects.

The idea of this approach is humans inspired; man can learn quickly by using what has been learned in the past. For example, a child can easily recognize the same person or animal in a large number of pictures.

Most approaches characterize few-shot learning as a meta-learning problem. To overcome the lack of data, a possible solution is to gain experience from other similar problems.

Meta-learning is a subfield of machine learning that is also known as learning to learn. In meta-learning, machine learning algorithms are applied to metadata related to machine learning experiments.

The main objective is to understand how machine learning can become flexible in solving learning problems and aims to improve the performance of existing learning algorithms or to learn (induce) the learning algorithm itself.

The small collection of labeled images used in meta-learning is called a support set as shown in Fig. 11. In contrast, the training set for conventional machine learning algorithms is large enough to learn a deep neural network, for example. Each class in the training set contains many samples.

The basic idea behind few-shot learning is that the support set has a limited number of classes and samples and can only provide additional information during the test. However, with a training set, if each class only has one sample, it is impossible to train a deep neural network.

The algorithm will be trained through a series of training tasks, each of which includes a support set with three different classes and two examples. The issue is a three-way-two-shot classification.

During training, the cost function will evaluate in turn for each task the performance on the query set taking into account the respective support set (Fig. 12). At test time, a completely different set of tasks is used to evaluate performance on the query set, given the support set.

As shown in Fig. 13, no overlap is existing between the classes in the two training tasks {cat, lamb, pig}, {dog, shark, lion} and between those in the test task {duck, dolphin, hen}. The algorithm will have to learn to classify image classes in general rather than focusing on classifying a particular set.

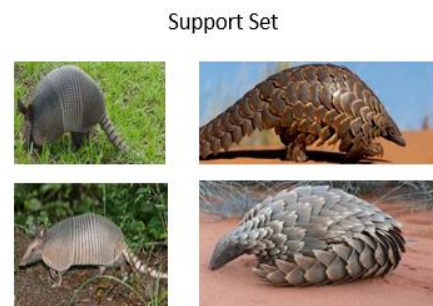


Fig. 11. Support Set Example.



Fig. 12. Query Set Example.

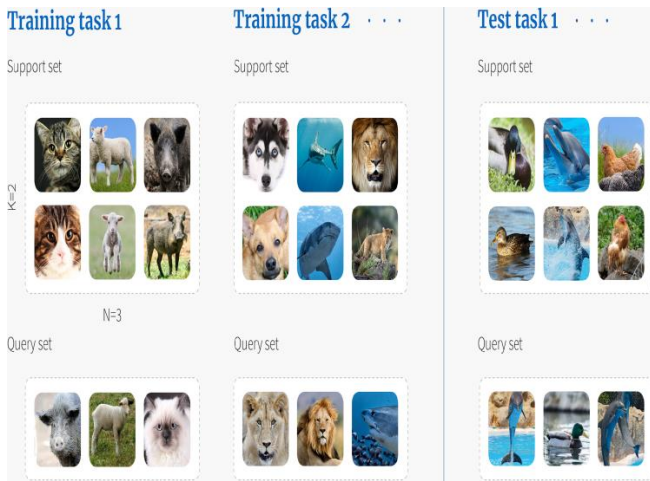


Fig. 13. Meta-Learning Framework.

The emphasis in FSL learning is on the quality of training data rather than the quantity. Furthermore, there is interest in designing and building AI machines or computer programs that improve automatically with experience.

The human-like learning allows FSL models to naturally advance robotics by improving robots capabilities that can at one-shot replicate or imitate human actions as well as enhancing their visual navigation.

- Zero-Shot Learning (ZSL)

Zero-shot learning is a learning model in which a machine is trained with an optimal minimum of labeled data during the learning phase. The machine learns to recognize a class of objects without having seen any previously labeled examples of that class. This method is also called on-the-fly learning.

Zero-shot learning relies on inference in order to reduce the requirement of the training phase for slightly different permutation masses.

The inference step in zero-shot learning is crucial: in this step, the algorithm attempts to predict and categorize classes of unseen data by analyzing its labeled data predictions to map the underlying attributes that have the highest probability of describing the data in general.

To solve Zero-Shot recognition problems, there are two popular ways. Fig. 14 depicts the anatomy of the first common method called “Embedding-Based Zero-Shot Learning”.

The input image is first processed by a feature extractor network (deep neural network (DNN)) to generate an N-dimensional feature vector for the image. This vector is fed into the main network, which produces a D-dimensional output vector.

The ultimate goal is to compute the weights of the projection network so that the N-dimensional input can be mapped to a D-dimensional output. Then, the loss compatibility module assesses the D-dimensional output's compatibility with the ground truth semantic attribute. The network's weights are tuned so that the D-dimensional output is as near as possible to the ground truth data.

The training seeks to develop a projection function from visual space to semantic space (word vectors or semantic embedding).

The Generative Model-Based Approach is the second Zero-shot Learning method. The generative method's goal is to use semantic attributes to generate image features for unobserved categories. At training time, the zero-shot classification model is trained on both seen and non-observed category images.

A general generative model-based zero-shot learning diagram is shown in Fig. 15.

The feature extractor network (deep neural network (DNN)) generates an N-dimensional feature vector for the image. The attribute vector is first fed into the generative model, as shown in the diagram. Based on the attribute vector, the generator creates an N-dimensional output vector. The generative model is trained in such a way that the synthesized feature vector resembles the original N-dimensional feature vector.

The generator's weights are fixed by the generative model, and the class attributes are used as input to generate non-observed category image features.

A basic image classifier is then trained by taking class image features (the training dataset) and non-observed category image features as the input and outputs the respective category label as shown in the Fig. 16.

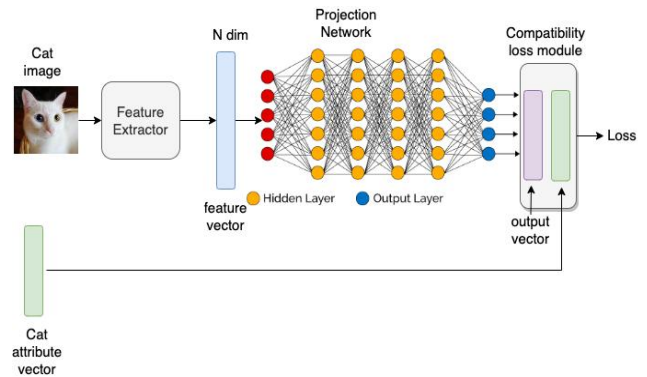


Fig. 14. Embedding-Based Zero-Shot Learning Model.

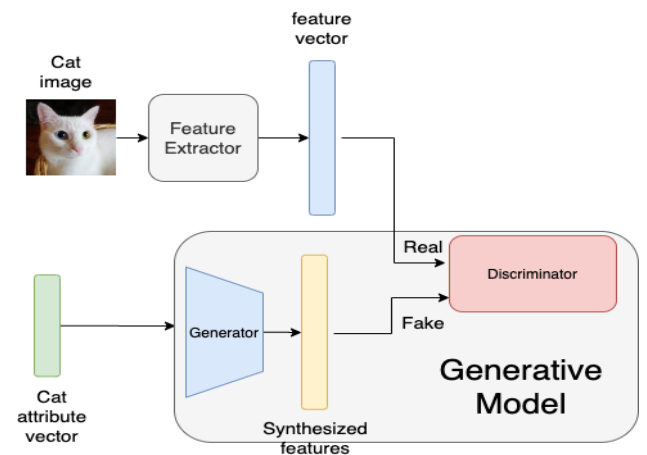


Fig. 15. Generative Model-Based Zero-Shot Learning.

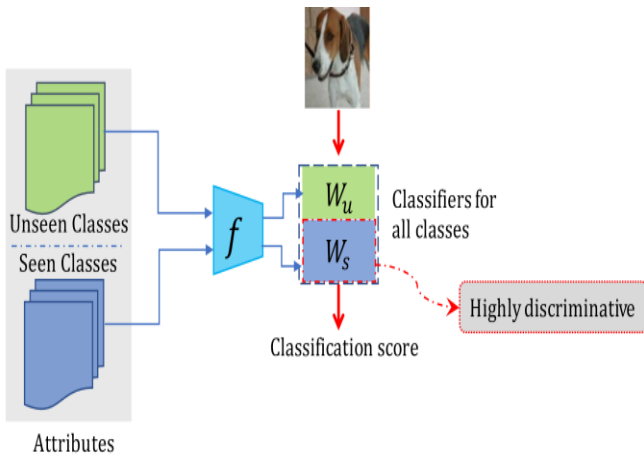


Fig. 16. Basic Image Classifier of Zero-Shot Learning Model.

**B. Proposed Model**

Our research aims to contribute to the growing literature on the new and novel theme that promises industrial revolution 5.0 in terms of collaboration between man and cobot. Our primary goal is to generate new ideas that will stimulate the interest of scholars in this subject.

According to our literature review, we have raised many challenges to overcome, especially when it comes to achieving effective collaboration between humans and machines. In this sense, the main problems that slow down the achievement of this goal largely concern big data.

Our approach consists of training a cobot by small amount of data using techniques discussed in this paper notably FSL and ZSL and MTDRL. An optimal common sense knowledge representation will be modeled and covers areas relating to human-machine collaboration in the context of Industry 5.0. These areas mainly concern actions and tasks, object recognition in the workplace and spatial navigation.

To compensate for the lack of the big data variable, research is continuing and focusing on strategies that rely simply on small data and allow for learning through collaboration or knowledge transfer.

In this order of thoughts, learning a cobot should be done as close to the human way as possible.

The cobot can collaborate and interact with its human counterpart only if it has a minimum of knowledge and common-sense background that allow it. The model will be able to learn at three levels: predictions against common-sense, human expectations and workers collaboration.

The cobot can act or accomplish a task in accordance with what has been pre-established in the model while remaining in permanent multimodal communication with the human collaborator who can correct it instantly to achieve the desired performance.

Like human learning, the cobot will be trained gradually from early stages of task completion, interaction, communication or decision making.

The model will be built around memory, computational units and three neural networks serving as training and correlating tools. The policies, rules and an optimal common-sense-knowledge repository will be stored at the memory level.

As shown in Fig. 17 our proposed model will be developed on three functions in order to calculate the action value, the policy and common-sense repository compliance value and finally the correlation function to assess the completion of multitasking based on the established policies and the repository of common-sense.

These functions will be implemented by three Deep Neural Network called:

- ZSL-DNN (Zero-Shot Learning Deep Neural Network).
- FSL-DNN (Few-Shot Learning Deep Neural Network).
- MTDRL-DNN (Multi-Task Deep Reinforcement Learning Deep Neural Network).

The Cobot's computational system will perform correlations between the objectives targeted by an action and what has been achieved by taking as reference the common-sense-knowledge repository, Human expectations and workers collaboration according to the algorithm illustrated in Fig. 18.

The algorithm's instructions will be carried out in the following sequence:

- Step 1: start of action;
- Step 2: correlation against common-sense;
- Step 3: return to step 2 if the outcome is unsatisfactory;
- Step 4: correlation against Human expectations;
- Step 5: return to step 4 if the outcome is unsatisfactory;
- Step 6: correlation against co-workers collaboration & correction;
- Step 7: return to step 6 if the outcome is unsatisfactory;
- Step 8: calculation of the action-value;
- Step 9: return to step 1 if the outcome is unsatisfactory;
- Step 10: reaction on the environment.

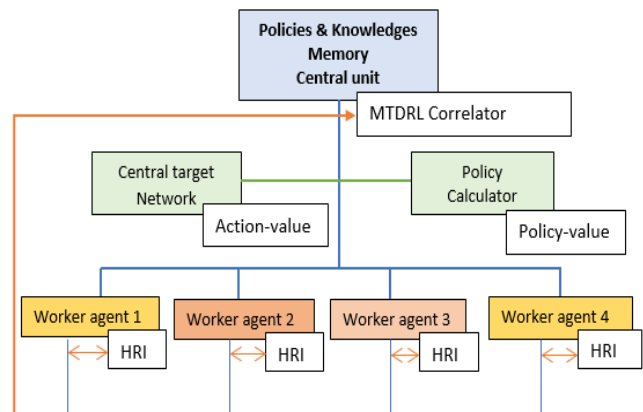


Fig. 17. Architecture of the Proposed Model.

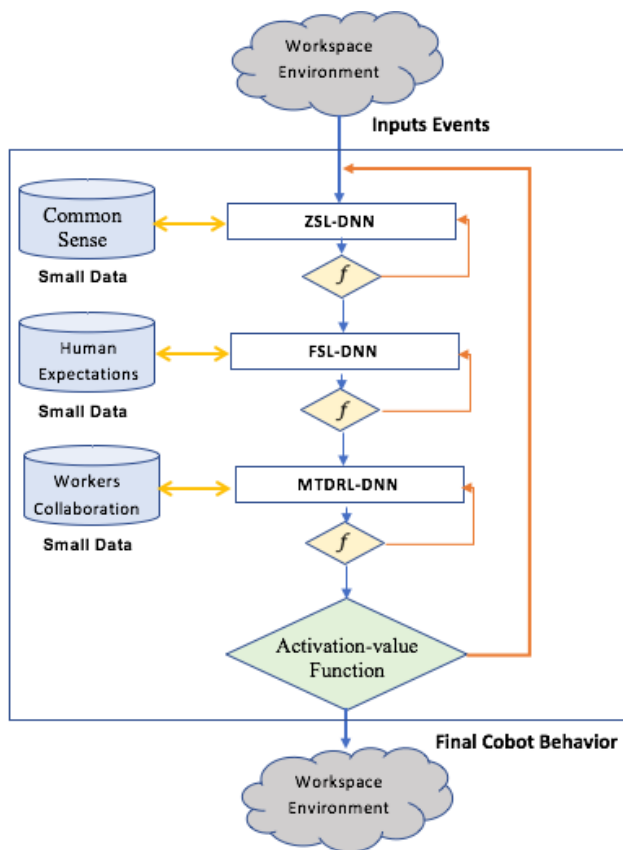


Fig. 18. Cobot Behavior Algorithm.

## V. CONCLUSION AND FUTURE WORK

The current trend in industry 5.0 is to be human-centered, the human and the cobot can interact safely and cooperate to accomplish the assigned tasks. The great progress of these systems is convincing, but the performance in this field is still far from being efficient. Artificial intelligence, while making great steps in many areas will have to overcome difficulties inherent to real-world environments.

Big data is critical to intelligence's success. On the other hand, the most significant constraint of deep learning is the requirement for enormous volumes of data; nevertheless, isn't it time to develop machines that can learn from little amounts of data?

In this paper, we have identified some of the technical challenges faced by researchers working on human-cobot collaboration. Current models of reinforcement learning for multi-tasking have many shortcomings that need to be addressed.

This study also highlights some of the existing solutions for addressing the key difficulties in the reinforcement area, such as DISTRAL, A3C, IMPALA, and PopArt.

Similarly, human-robot interaction has been briefly discussed given its importance in achieving Industry 5.0 goals. Human-robot interaction implies many technical challenges both in technical and human-centered aspects. Human-robot interaction is an open multidisciplinary field where current research is alive and growing.

Finally, we have proposed a new model and algorithm for training Cobots from small amount of data. Our model is based on three Deep Learning Neural Networks such as ZSL, FSL and MTDRL. The Cobot training will be gradually from early stages of task completion, interaction, communication or decision making.

In future work, we will try to deploy and experiment our model in a real industrial 5.0 context, in order to establish an appropriate Cobot collaboration.

## REFERENCES

- [1] M. Peshkin and J. Colgate, "Cobots," *Industrial Robot*, An International Journal, vol.26 n°5, pp. 335-341, 1999.
- [2] B. Claverie, B. Le Blanc and P. Fouillat, "La Cobotique," Presses universitaires de Bordeaux, 1 décembre 2013, DOI: 10.4000/communicationorganisation.4425.
- [3] Christoph. Bartneck, "A design-centred framework for social human-robot interaction," *University of Canterbury Conference Paper* October 2004.
- [4] Kristinn R.Thórisson, Jordi.Bieger, Thröstur.Thorarensen, Jona S. Sigurdardottir, "Why Artificial Intelligence Needs a Task Theory – And What It Might Look Like," *Conference Paper* July 2016, DOI: 10.1007/978-3-319-41649-6\_12.
- [5] Francesco. Longo, Antonio. Padovano and Steven. Umbrello, "Value-Oriented and Ethical Technology Engineering in Industry 5.0: A Human-Centric Perspective for the Design of the Factory of the Future," *applied sciences*, 18 June 2020.
- [6] Zhikun. Wang, Katharina. Mülling, Marc Peter. Deisenroth, Heni. Ben Amor, David. Vogt, Bernhard. Schölkopf, Jan. Peters, "Probabilistic Movement Modeling for Intention Inference in Human-Robot Interaction," Vol. 32, issue. 7, page(s). 841-858, *SAGE Journals*, April 2013.
- [7] Pieter. Abbeel and Andrew Y. Ng, "Apprenticeship Learning via Inverse Reinforcement Learning," *Computer Science Department, Stanford University: Stanford, CA 94305*, September 2004, DOI:10.1007/978-0-387-30164-8\_417.
- [8] Manuel. Lopes. Francisco S. Melo, Luis. Montesano, "Active Learning for Reward Estimation in Inverse Reinforcement Learning," *Conference: Machine Learning and Knowledge Discovery in Databases, European Conference, ECML PKDD 2009, Bled, Slovenia, September 7-11, 2009, Proceedings, Part II, September 2009*, DOI:10.1007/978-3-642-04174-7\_3.
- [9] Zhiyuan. Xu, Kun. Wu, Jian. Tang, Jieping. Ye, "Knowledge Transfer in Multi-Task Deep Reinforcement Learning for Continuous Control," *Conference Paper: 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada, October 2020*.
- [10] John. McCarthy, "WHAT IS COMMON SENSE," *Computer Science Department, Stanford University : Stanford, CA 94305*, February 2002.
- [11] Antonio. Bicchi, Michael A. Peshkin, J. Edward Colgate, "Safety for Physical Human-Robot Interaction," *Springer Handbook of Robotics*, pp.1335-1348, Springer, Berlin, anuary 2008, DOI:10.1007/978-3-540-30301-5\_58.
- [12] Jean. Scholtz, "Theory and Evaluation of Human Robot Interactions," *Conference: 36th Hawaii International Conference on System Sciences (HICSS'03), January 2003*, DOI:10.1109/HICSS.2003.1174284.
- [13] Shirine. El Zaatari, Mohamed. Marei, Weidong. Li, Zahid. Usman, "Cobot programming for collaborative industrial tasks: An overview. Robotics and Autonomous Systems," *Publisher: Elsevier, Jun 2019*, DOI 10.1016/j.robot.2019.03.003.
- [14] Murtua, I., Fernandez, I., Tellaeché, A., Kildal, J., Susperregi, L., Ibarguren, A., Sierra, "Natural multimodal communication for human-robot collaboration. *International Journal of Advanced Robotic Systems*," July-August 2017, DOI: 10.1177/1729881417716043.
- [15] Ding. Z, Dong, H, "Challenges of Reinforcement Learning In Deep Reinforcement Learning," *Springer: Berlin/Heidelberg, Germany, 2020*.
- [16] Ruben. Glatt, Anna. Helena Reali. Costa, "Improving deep reinforcement learning with knowledge transfer," *In Proceedings of the*

- Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
- [17] Hessel. M, Soyer. H, Espeholt. L, Czarniecki. W, Schmitt. S, van Hasselt. H, “Multi-task deep reinforcement learning with popart,” In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019.
- [18] Matthew E. Taylor, Peter. Stone, “An introduction to intertask transfer for reinforcement learning,” *Ai Magazine* 32(1), March 2011, DOI:10.1609/aimag.v32i1.2329.
- [19] M. Betke, W. Mullally, and J. Magee, “Active detection of eye scleras in real time,” In Proceedings of the IEEE Workshop on Human Modeling, October 2001.
- [20] M. Betke, W. Mullally, and J. Magee, “Active detection of eye scleras in real time,” In Proceedings of the IEEE Workshop on Human Modeling, October 2001.
- [21] E. Horvitz and T. Paek, “Harnessing models of users’ goals to mediate clarification dialog in spoken language systems,” In Proc. of the Eighth International Conference on User Modeling, 2001.
- [22] Dagen. Wang and Shrikanth Narayanan, “An acoustic measure for word prominence in spontaneous speech,” *IEEE Transactions on Speech, Audio and Language Processing*, 15(2):690–701, Feb 2007, DOI:10.1109/TASL.2006.881703.
- [23] Rizzolatti. G, and Arbib. M.A, “Language within our grasp,” *Trends in Neurosciences*, 21(5):188–194, 1998.
- [24] B. Scassellati, “Investigating models of social development using a humanoid robot,” *Proceedings of the International Joint Conference on Neural Networks*, 4:2704–2709, July 2003.
- [25] C. Breazeal, A. Edsinger, P. Fitzpatrick, and B. Scassellati, “Active vision for sociable robots,” *IEEE Transactions on Systems Man and Cybernetics - Part A Systems and Humans* 31(5):443 – 453, DOI:10.1109/3468.952718.
- [26] Nelson Vithayathil. Varghese and Qusay H. Mahmoud, “A Survey of Multi-Task Deep Reinforcement Learning,” *Electronics* 9(9):1363, August 2020, DOI:10.3390/electronics9091363.
- [27] Yongyuan. Liang, Bangwei. Li, “Parallel Knowledge Transfer in Multi-Agent Reinforcement Learning,” arXiv, arXiv:2003.13085, March 2020.
- [28] Yee Whye. Teh, Bapst. V, Czarniecki . M, Quan. J, Kirkpatrick. J, Hadsell. R.; Heess. N, Pascanu. R, “Distral: Robust multitask reinforcement learning,” In Proceedings of the Thirty-first Annual Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9, December 2017; pp. 4496–4506.
- [29] Liu. X, Li. L, Hsieh. P.C, Xie. M, Ge. Y, “Developing Multi-Task Recommendations with Long-Term Rewards via Policy Distilled Reinforcement Learning,” arXiv 2020, arXiv:2001.09595, January 2020.
- [30] Andrei A. Rusu, Sergio Gomez. Colmenarejo, Caglar. Gulcehre, Guillaume. Desjardins, James. Kirkpatrick, Razvan. Pascanu, Volodymyr. Mnih, Koray. Kavukcuoglu, and Raia. Hadsell, “Policy distillation,” *International Conference on Learning Representations (ICLR)*, 2016.
- [31] Emilio. Parisotto, Jimmy Lei. Ba, and Ruslan. Salakhutdinov, “Actor-mimic: Deep multitask and transfer reinforcement learning,” *International Conference on Learning Representations (ICLR)*, 2016.
- [32] Mnih. V, Badia. A.P, Mirza. M, Grave. A, Lillicrap. T, Harley. T, Silver. D, Kavukcuoglu. K, “Asynchronous methods for deep reinforcement learning,” In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016.
- [33] Wang. Y, Stokes. J, Marinescu. M, “Actor Critic Deep Reinforcement Learning for Neural Malware Control,” In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12; Volume 34, pp. 1005–1012, February 2020.
- [34] Espeholt. L, Soyer. H, Munos. R., Simonyan. K, Mnih. V, Ward. T, Doron. Y, Firoiu. V, Harley. T, Dunning. I, “Impala: Scalable distributed deep-rl with importance weighted actor-learner architectures,” ArXiv. Feb 2018.