World Journal of Advanced Engineering Technology and Sciences

eISSN: 2582-8266

Cross Ref DOI: 10.30574/wjaets

Journal homepage: https://wjaets.com/

(Research Article)

Check for updates

Human-in-the-Loop Robotic Process Automation (RPA) with Deep Reinforcement Learning (DRL)

Abhaykumar Dalsaniya *

Principal Automation Architect, USA.

World Journal of Advanced Engineering Technology and Sciences, 2020, 01(01), 060–076

Publication history: Received on 14 October 2020; revised on 23 November 2020; accepted on 27 November 2020

Article DOI: https://doi.org/10.30574/wjaets.2020.1.1.0015

Abstract

In this article, the authors attempt to extend human feedback as a critical decision-making tool to improve DRL for RPA systems, especially in application domains sensitive to change, such as healthcare and legal practice. The previous methods of implementing RPA moved in well-defined trajectories, meaning they have low adaptability for complex decision-making or other atypical cases. These systems can integrate DRL and hence can learn and evolve over the period required. Nevertheless, DRL models often work independently from human beings. They are used in applications with low human interaction, which can cause issues in decision-making procedures that require a higher understanding of regulation.

The article puts human-in-the-loop, with human experts returning to the DRL models. It encapsulates human feedback into meaningful rewards or penalties for the DRL algorithms, ensuring correction and improvement of decision-making brought about by the expert's input. This approach can enhance the management of exceptions/corner cases typical to regulated domains, the automation of which must consider compliance and legal requirements.

The paper focuses on how human intervention can be combined in DRL to augment the reward schemes that are superior to those provided by purely automated systems to human overseers. The efficiency of the hybrid system of DRL coupled with RPA in real-world use cases that involve processing legal documents and medical records is well supported. Currently, feedback from people in DRL for RPA has the potential to create more dynamic, dependable, and adherent robotic processes in various stringent compliance contexts.

Keywords: DRL Models; Machine Learning; Deep Reinforcement Learning; RPA; Robotic Process Automation; Interactive Machine Learning

1. Introduction

RPA, also known as Robotic Automation, is a new technology that automates many repetitive work fields, permitting different organizations to become more efficient and save money. RPA broadly means technology solutions where machines replicate human actions to perform repetitive tasks, including inputting data, processing invoices, and taking other functions. It allows companies to make work more efficient by eliminating the need for manual input and decreasing the likelihood of mistakes. Dramatic results have been observed in banking, insurance, and telecommunications, where repetitive work is predominant (Lacity, Willcocks, & Craig, 2015). For example, at Telefónica O2, RPA was used to automate back-office work, which enabled large operational gains in cost-cutting and higher service delivery (Lacity, Willcocks, & Craig, 2015).

* Corresponding author: Abhaykumar Dalsaniya.

Nevertheless, the traditional RPA systems have two main drawbacks: they do not allow the incorporation of knowledge-intensive activities into automation processes. These systems are based on predefined checklists and standard procedures of how people should work and, therefore, need to be better suited to highly changeable business contexts where decisions can and must be made ad-hoc from the continuously updated data. This is where the potential Deep Reinforcement learning (DRL) comes into play. Deep reinforcement learning, or DRL, is a subfield of machine learning that incorporates deep learning with reinforcement learning to allow a system to determine a decision-making process by experiencing its environment. DRL makes machines learn better actions at given steps, and with the help of feedback, the next actions are formed. It has achieved human-like ability in such applications as video game playing and mastering game environments to multiply automation (Mnih et al., 2015).

When combined with RPA, DRL could increase the sophistication of automated systems, especially where business processes and compliance are complex. DRL helps RPA systems break free from conventional rule-governed working and work smartly by continually learning the next course of action on their own. For instance, when the function of a DRL-based RPA system applies to legal or medical document processing, such a system could learn from previous experiences and iterations of the past and generate improved results in the end (Mnih et al., 2015). Nevertheless, using autonomy in DRL systems creates the feeling of safety, reliability, and conformity with the laws and regulations. For applications of DRL in fields like healthcare or law, the cost of a mistake is significantly costly; it is, therefore, imperative that the models can make industry-appropriate decisions consistent with best practices, standards of training, and the law.

To overcome these challenges, there are recent approaches known as Human-in-the-Loop (HITL) systems. HITL systems entail using human beings to review and supervise the results returned by artificial intelligence to conform to the ideal human-generated results with models that rely on predisposed human judgments and industry practices. Instead, interactive machine learning is defined by the incorporation of inputs from the human beings in the correction of the model used to make the decisions as well as in increasing the efficiency of the decision making as well as the accuracy levels (Interactive Machine Learning, 2014). This is quite mind-blogging that with human input, DRL models can be steered toward a better choice specifically in cases that may need professional assessment. In professions with many rules and regulations, like healthcare or law, the HITL system guarantees that automated operations are legal and ethical, minimizing costly mistakes.

To reintroduce human input into DRL models, human feedback should be translated into consequentialist rewards or penalties that are necessary by the model during learning. In this way, the potential of DRL-based RPA systems for developing the ability to learn to adapt increases, and the decisions made by these systems coincide with the development of mankind. For example, in a medical document processing system, a human expert could tag exceptions or anomalies – the DRL model could then take these tags as feedback for improved behavior (Amershi et al., 2014). It has huge prospects to enhance the reliability and safety of the RPA systems in the sectors that require specific guidelines because of the nature of their business to make the RPA systems more flexible and human-like.

1.1. Overview

Incorporating human feedback into Deep Reinforcement Learning (DRL) is an excellent enhancement for improving the performance of Robotic Process Automation (RPA) systems, especially for managing complex and dynamic tasks. DRL models normally involve a reward function with decisions based on accumulative rewards regarding a particular action. However, creating a reward function that will embody all aspects of the real world can be quite complicated, especially for organizations in a highly defined industry that requires precise decision-making and compliance with set rules and regulations. Human feedback offers a solution since DRL models can be advised to perform like humans who set the norms and follow legal requirements (Christiano et al., 2017).

Human feedback can be employed within the context of RPA to improve the DRL models in utilitarian sectors such as healthcare and law due to their highly stringent regulatory environments and the need to adhere to particular ethical guidelines. For example, human experts can explain exceptions or anything beyond which traditional business rules can be applied. This feedback can then be rewarded or punished, directing the DRL model towards decisions that increase its chances of high performance and adherence to the legal requirements and rules (Christiano et al., 2017).

One approach for teaching DRL agents is TAMER (Training an Agent Manually via Evaluative Reinforcement). This technique, he said, has been valid in enabling agents to change their actions in response to human ratings, enhancing the learning process, and accommodating real-life dynamics (Knox & Stone, 2009).

Moreover, when integrated, human feedback enables learning to proceed faster in large and especially complex systems. Suay and Chernova (2011) proved that human-assisted reinforcement learning not only increases the learning process's speed but also increases the accuracy of optimal decisions when the working environment is characterized by a massive number of states or complex working processes, which are typical for RPA systems.

1.2. Problem Statement

Applying Deep Reinforcement Learning (DRL) in regulated industries, including healthcare, finance, and legal services, is not without unique amplified challenges. They include sectors quite sensitive to legal compliance and tend to be closely regulated, with high compliance standards, ethical issues, and risk management issues. Challenges of integrating DRL in these industries stem from the fact that while DRL is highly effective, there needs to be more clarity in interpreting, auditing, and regulating how it arrives at its decisions, making it difficult to guarantee compliance with current laws.

DRL has been experiencing difficulty in the regulated industry where matters of law and ethics come into play due to the automated systems' policy limitation in resultant actions. DRL models learn with the help of a trial and error method while maximizing future gain. However, in many industries, the strategies that are optimal from an algorithm's perspective may violate a regulation or an ethical consideration. This holds a lot of dangers, especially when applied to sensitive areas such as health, where non-adherence can cause harm to patients or finances, where a mistake can land the company in legal trouble or even make it financially immobile (Amodei et al., 2016).

The next challenge relates to the openness and interpretability of the current DRL systems. It is recorded that regulated industries demand control, and an automated system's decision must be well explained or justified. Deep learning models, particularly DRL ones due to their uninterpretability, are known as "black boxes." Due to this lack of transparency, regulators cannot apply their confidence and permission for the use of DRL-based systems, as they have to ensure that these systems adhere to guidelines, ethical principles, and regulations (Dulac-Arnold, Mankowitz & Hester, 2019).

Third, the nature of real-world environments by default makes it a challenging task to implement DRL systems, especially in regulated industries. Such industries can be faced with exceptional cases or edge circumstances to be resolved that cannot be described with criteria or profiles. DRL performs a good job when the environment is complex, but it still needs improvement, and it can be wrong sometimes in conditions that include significant but sparse events. This volatility makes non-compliance or error occurrence more probable as the worst among regulated industries.

1.3. Objectives

The main objectives of this research are to:

- Incorporate human feedback into DRL algorithms: Find methods that add human factors into the system to help DRL models toward better decision-making.
- Enhance decision-making in RPA systems for regulated industries: Enhance the existing RPA to satisfy better compliance and decision-making requirements of sectors such as healthcare and finance.
- Increase interpretability of DRL-based RPA systems: Build comprehensible models, which can help solve the problem of weak regulation.
- Adapt RPA systems to handle exceptions: Enhance the practical capacity of RPA systems to handle special tedious cases by evaluating the feedback of trainers and data collected in real time.
- Ensure ethical and regulatory compliance: Emphasize the safety of the implemented RPA systems and make it mandatory to meet the required regulatory standards so that its implementation in sensitive areas will not pose excessive risks.

1.4. Scope and Significance

This research topic uses human feedback to augment Deep Reinforcement Learning (DRL) in Industrial Robotic Process Automation (RPA) in sectors requiring high regulatory compliance, including legal and medical fields. The motive of the proposed approach is to overcome the limitations of the conventional RPA systems where the process tasks are based on a predefined set of patterns and are incapable of dealing with the cases based on exceptions. Thus, with the help of human input, the complexities of such scenarios can be managed, and DRL models can be used to make the RPA systems not only smarter but also capable of handling sophisticated, high-risk decision-making. At present, in medicine, the proposed DRL-based RPA system with human feedback integration would be useful in improving specific tasks like medical records processing or diagnostic aid or following data protection laws related to patients' information. These systems can also train from raw input human data to improve decision steps, making them less inclined to produce wrong results that can affect the patients. In the legal profession, the RPA systems, when combined with DRL, can help with document quality review, legal research, and case analysis while making decisions, the results of which must follow specific rules and initiatives to that effect have been introduced. Such systems require help from humans in dealing with special legal cases and loose situations that cannot be formalized in rules. Several limitations can be erroneously perceived as benefits while utilizing such a strategy. The main issue comes from feedback from humans and having them directly and easily correlate to the reward signals for the DRL models. Hence, making such systems ‘explainable’ to meet very high legal and ethical requirements is very important. However, the opportunities to enhance smarter, more flexible, and compliant RPA solutions in the regulated sectors render the current study valuable for developing the automation system.

2. Literature review

2.1. Robotic Process Automation (RPA)

One of the most emerging technologies in the current business environment is Robotic Process Automation (RPA) since it can automate repetitive jobs because they depend on explicit rules. Robotic Process Automation (RPA) automates business processes through software robots that imitate human actions in performing basic clerical tasks like data entry, invoice processing, customer relations, etc. With the help of RPA, different kinds of costs have been decreased, productivity has been improved, and human mistakes in business processes have been minimized (Willcocks, Lacity, & Craig, 2015).

RPA technologies are widely employed in finance, health care, telecommunications, and insurance sectors. Such sectors are generally characterized by numerous repetitive processes that may easily be streamlined to reduce costs. For example, using RPA in financial services includes account reconciliation, invoice processing, and regulatory reporting processes. It enabled banks or financial institutions to have high accuracy with less pressure on the employees. In other words, mass automation cuts a lot of employees' workloads (Willcocks et al., 2015).

RPA is used in record management, appointment setting, and insurance claims handling as a primary application in healthcare organizations. It also relieves the pressure of charging and billing processes on healthcare givers and greatly improves patient data's legal and efficient handling. Outsourcing of back-office tasks means that healthcare providers can devote more time to actual patient care, improving total healthcare (Aguirre & Rodriguez, 2017).

Many telecom businesses also reap from RPA, particularly for customer support and networks. First, RPA bots can process and resemble many customer inquiries and requests; second, they can quickly identify and solve network problems, enhancing service delivery and customer satisfaction. This has made it easy for the telecom industries to use RPA to expand their business while at the same time offering quality customer support services, as noted by Willcocks et al., 2015.

As suggested above, RPA has its disadvantages. Most RPA systems are coded and work based on rules hence each RPA system can address only routine processes that are more or less standardized. In information that is not easily divided into categories and processing tasks where discretion and initiative are required, it has numerous issues.

However, RPA systems have some limitations, such as the system's inability to adapt to constantly changing environment contexts, which always need to be calibrated (Lacity & Willcocks, 2016). Other limitations arise, requiring more complex technologies to complement RPA systems, including artificial intelligence and machine learning, as relevant industries grow over time.

Further evolution of RPA systems is anticipated to gain from incorporating artificial intelligence (AI) and machine learning (ML) into the software. This development would allow the RPA bots to perform more sophisticated data-intensive operations where the bot would employ lessons learned from previous or present interaction scripts with the client system. Heralded as intelligent automation, these systems can be integrated at an industrial scale, built to extend traditional RPA while incorporating cognitive technologies that enable large-scale improvement in business processes (Lacity et al., 2015).

2.2. Deep Reinforcement Learning (DRL)

DRL stands for Deep Reinforcement Learning, which integrates deep learning and reinforcement learning to solve decision-making problems. In the theory of RL, an agent has to make the right move (mapping into environment states) to achieve the highest total reward possible over some time. The agent uses the information the environment provides by either a positive or negative signal and revises the information accordingly. Nevertheless, RL has challenges handling high-dimensional input space, such as image inputs, making it irrelevant for most practical real-world problems involving complex inputs. RL is improved by deep learning since it processes input data using the neural network, making it suitable for such forms of data (Li, 2018).

Hence, the basic idea in DRL is to use deep neural networks to estimate the value functions and policies that govern the behavior of an agent in an RL setting. In this setting, an agent identifies the state of the environment, takes an action according to the learned policy, and gets its rewards or penalties. As time progresses, the agent knows the optimal policy to maximize the discounted reward and often employs techniques such as Deep Q-Learning (DQN), policy gradients, or action-value methods such as the Actor-Critic. Due to the efficient imitation of high-dimensional functions, DRL has recorded impressive performance in game AI, robotics, and self-driving cars (Li, 2018).

DRL has two main successes, illustrated below: The creation of Deep Q-Networks (DQN), which embedded utilities deep into agents and made them play games better than humans. In late April 2015, a DeepMind team showed that DRL can dominate games such as Atari's Breakout or Space Invaders without being tweaked for the particular task. The DQN algorithm used the deep neural network to estimate the Q-value to directly train the agent from pixel inputs (Mnih et al., 2015). This paradigm shift in AI demonstrated how DRL could solve large decision-making issues with little help from human experts.

However, apart from games, DRL has been promising in robotics engineering. With the help of DRL, robots are trained to acquire complex environment information, manipulate objects, and interact with the surrounding environment more actively. This profoundly affects areas like production and health, where DRL can handle boosted robots. These complex operations demand hand work, decision-making, and more from human workers (Levine et al., 2016). Thus, in DRL, robotics yields enhanced performance in given environments, thereby preventing the need to separately code enhancements for subsequent choreography in ROB.

The sphere of autonomous vehicles was another area that DRL actively fostered throughout the period. DRLs have also been used in the training of self-driven vehicles regarding matters such as objects and ways, routes to choose, and real-time decisions. As a result, such vehicles can adapt their operation because of their environment; they can prevent traffic accidents and facilitate traffic flow. It is especially used in this self-driving car to coordinate future transportation directions (Kendall et al. 2018).

But, these are real-world success stories, and several issues have been encountered in implementing DRL in real-world problems. DRL depends on heavy data and computational facilities; the time spent on training can be exhaustive and expensive. However, it is sometimes challenging to understand DRL systems because the workings of a deep neural network frequently need to be clarified. The above lack of transparency raises concerns in such industries because decision explainability is crucial (Li, 2018).

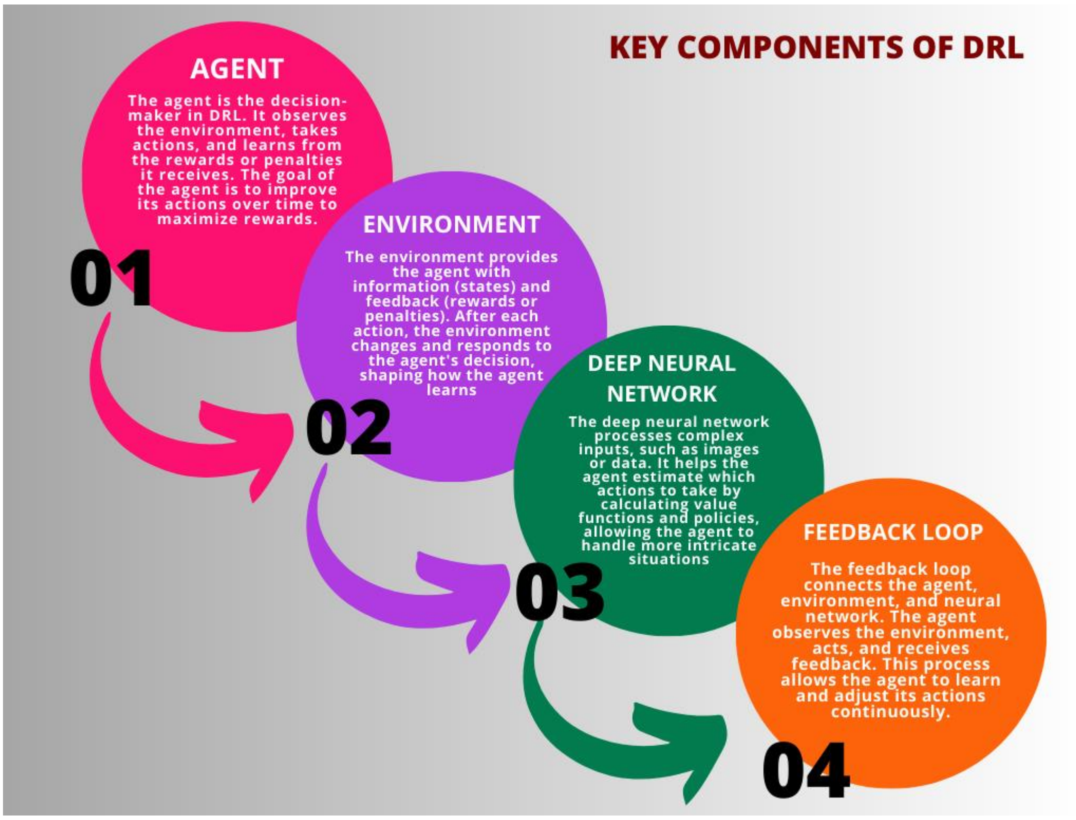

Figure 1 Image showing the Key Components of DRL

2.3. Human-in-the-Loop Systems

Human-in-the-loop (HITL) systems tend to include human inputs during the process of model learning, thereby improving on the results given by the algorithmsIn typical machine learning architectures, solutions are learned through a set of data samples without active intervention from humans in real-time, and this makes decision-making for systems to be suboptimal whenever data is noisy or ambiguous. The marriage of AI and human beings provides a solution in that models can learn from iterative input from humans and provide more refined outputs that allow the incorporation of expert knowledge in subsequent iterations (Fails & Olsen, 2003).

HITL or Human-In-The-Loop refers to the subdomain of the above cases focusing on non-stop interaction between humans and the result-generating ML model. In IML, human users can then interfere and decide on something during data labeling, correction, or regular during exaction of the model. In addition, such interaction enables the machine learning models to learn more with fewer data points, making the systems more flexible and adequate in new, complicated, or obscure situations. The advantages of human intervention are even more significant, as automated systems underperform in handling the outliers or exceptions that the formalisms cannot easily accommodate (Fails and Olsen 2003).

Another of the primary uses of human interaction in IML is to specify that the system incorporates expert feedback capability to learn from the feedback in real-time. In such intricate areas as, for example, health or interpretation of legal papers, one can get a second opinion from a human concerning the system's choices and fine-tune the algorithm. For instance, in a legal RPA system based on deep reinforcement learning (DRL), a legal expert can intervene and give feedback on how the system should handle cases that were earlier labeled as having equivocal management by the system to enable the model to learn improved decision rules. In the long run, the system becomes enhanced with the capability to integrate expert judgments when making its own automated decisions, as posited by Amershi and colleagues in the study.

In HITL systems, how the systems operate has been particularly effective in different fields. For instance, HITL has been used in robotics to assist a robot in mastering various tasks by demonstration and reinforcement. The robots may thus learn more from human users during fewer learning or training sessions, which is preferable in scenarios where large data sets are difficult to obtain (Suay & Chernova, 2011). Likewise, in the image classification task, the correction made by the human user concerning the mistakes made by the learning system converges faster. It gives better results than conventional machine learning methods (Fails & Olsen, 2003).

However, HITL systems are fine. One of the main challenges is the issue of the amount of human intervention possible. This means that if the AI depends too much on human input, the learning rate will be slow, and if a human interacts with the AI very little, the training results will be incomplete or even wrong. However, HITL applications require effective and natural interfaces that do not hinder the human-endogenous system interaction (Amershi et al., 2014). Substantial HITL systems must utilize fully automated and partially supervised systems to attain the highest performance.

2.4. Human Feedback in DRL

Using human observation as a feed-forward to the networking learning algorithms, including the DRLhas, is now a crucial approach to improving the choosers for the AFs. Most DRLs rely on the use of the environment by the agent for rewards or penalty outcomes from the actions that are executed. Authorizing teaching feedback for complicated real-life assignments can be demanding, even where the tasks have aspects of ethical or legal requirements. I addressed this challenge by applying human preferences within the DRL models and enabled the systems to learn from human expertise independently (Warnell et al., 2018).

One way human feedback has been incorporated into DRL is called Deep TAMER, which is intended to train agents in high-dimensional states using human input (Warnell et al., 2018). In Deep TAMER, unlike traditional TAMER, an actual human being watches an agent as it learns and adjusts agent behaviors concerning its learned preference function independent of the reward function. This makes DRL models capable of learning from fewer endeavors than possible with purely conventional methods and makes them suitable suits for tasks where human discretion is an essential commodity.

However, Deep TAMER is especially useful when it is complicated to state what precisely needs to be achieved and when the reward function is vague, for example, when solving an ethical problem or when the task relies on intuition. For instance, in developing a medical decision-making system, human feedback can help to direct the flow of DRL to arrive at better choices depending on experience and thus knowledge and not just the whole reward. This technique puts into consideration the push and pull factors of life and the ethical working standard of the professional society (Warnell et al., 2018).

Another way to incorporate human feedback into DRL is using Reinforcement Learning from preferences, commonly referred to! Here, human evaluators' exploitation of the policy involves ranking one action or outcome against another and providing feedback for adaptation of the policy. In a study by Christiano et al. (2017), it became clear that human preference-based reinforcement learning would also boost the performances of DRL systems such as video game-playing systems. Feedback collected in the form of human preferences then enables the DRL agent to understand which of its actions led towards the accomplishment of goals that are more in line with the human preferences even when the reward function is not very well defined or comprehensible by a machine (Christiano et al., 2017).

Caring for human preferences also directly solves the familiar exploration-exploitation problem in reinforcement learning. In DRL, at this step, an agent decides whether to solve a new action space or utilize the described corridors, contributing to a high reward. Random action generation can be avoided with human feedback providing the information on which actions are worthy of being taken, thus expanding the learning process (Zhang et al., 2019). This results in better learning, as comprehensively explained below, especially when working with large and continuing state spaces, which may slow down exploration.

However, like other approaches, integrating human feedback into DRL also has some concerns. One concern is dispersing attention to human evaluators, who always have to give feedback on the learning process. This can be prevented by organizing feedback-asking systems only to require input in certain circumstances or, most importantly, designing feedback-requiring systems that are not complex (Warnell et al., 2018). However, human input in such methods must be completely uniform and free from biases that could downgrade the effectiveness of such methods because inconsistent feedback normally avails suboptimal learning effects.

2.5. RPA in Regulated Industries

One of the emerging digital technologies that enjoys more usage in various industries is Robotic Process Automation (RPA), especially in the healthcare and legal services sector, as it causes maximal impact on effectiveness, accuracy, and compliance. These industries deal with organizations' data and critical processes, which cannot afford to be a subject of regulatory compliance. Applying RPA in these sectors optimizes the processes and simultaneously keeps high compliance and data security indicators (Asatiani & Penttinen, 2016).

On the healthcare side, the function of RPA is used to ease processes such as arranging appointments, invoicing, handling of claims, and managing health information. By automating these activities, healthcare financial managers can cut operational costs, enhance coordination accuracy, and have more medical staff invest time in patients. For example, RPA can help organizations reduce insurance claim processing time by extracting and validating data from the forms where such claim is submitted (KPMG, 2017). Further, as one of the applications, it contributes to the proper record of patient data, legally known as electronic health records, by updating patients' information and appointments (Willcocks & Lacity, 2016).

However, some issues are associated with applying RPA in the healthcare sector. The industry is highly regulated; for example, the Health Insurance Portability and Accountability Act (HIPAA) has policies and measures that must be adhered to regarding data management and protection. Due to such regulations, RPA systems must be developed to meet them; security measures that will enable them to protect patient information are implemented. If this is not done, the consequences are legal implications and loss of trust and confidence from the patient.

RPA can be used for document review, research, and compliance in the legal profession. There is always a lot of documentation dealt with in law firms, which needs to be analyzed and likely classified. This can be done by automating the review process through RPA bots and, by doing so, extracting any important data from any documentation. Attorneys could save time and focus on more analytical tasks (Antunes & Gill, 2018). For instance, using RPA, contracts can be read and reviewed to check for clauses or compliance with the required legal standards; this eliminates cumbersome manual checking, and the chance of human error is dismissed (Reade, 2016).

During M&A transactions, RPA helps evaluate large documents containing certain risks and obligations and makes the results quicker. This expedites the process of doing due diligence and improves the effectiveness of the results (Antunes & Gill, 2018). However, like the healthcare system, client requirements for RPA solutions must meet data privacy laws within the legal sector. RPA systems should be set up correctly to avoid legal breaches and address sensitive data.

One of the best examples of the positive effects of RPA in industries strictly regulated is the OpusCapita company, focusing on financial processes. RPA was implemented in OpusCapita by integrating it to reduce the time taken to complete the processes while at the same time leading to cost savings and a highly improved method of delivering services. More transactions were processed, and fewer mistakes were made, which reveals how RPA can be applied adequately in industries with remarkable compliance constraints (Asatiani and Penttinen, 2016).

2.6. Challenges in DRL for RPA

Real-world applications of Deep Reinforcement Learning (DRL)-based Robotic Process Automation (RPA) systems pose a major safety and reliability challenge. A major challenge is that the environment specifying an optimal course of action often differs vastly from the surroundings in which DRL models are trained. While in environments such as gaming, agents do not face the negative real-life implications of unsafe choices, the same is not always the case in raw operating environments; decisions may have catastrophic consequences, particularly in heavily supervised industries (Dulac-Arnold et al., 2019).

One of the most significant drawbacks of DRL algorithms is the sample size used in the algorithms' performance. DRL models take millions of iterations to realize effective policies in their surrounding environment. In practical, real-world RPA systems such as the above, exploration is not feasible because of time issues and the costs of errors. This inefficiency slows down the use of DRL in settings where data acquisition is costly and or dangerous (Dulac-Arnold et al., 2019). For instance, in financial transaction processing, an incorrect action could cause serious monetary loss and violation of the set law.

One of the issues is the safety when exploring. There are new learning mechanisms in DRL agents that allow agents to learn from different actions, including non-safe or non-adherent actions related to industry standards. The environment might be hazardous, or an agent needs the ability to watermark large areas or navigate while it's not safe for the user. There is significant uncertainty about whether the agent's learning process must be secure and how not to endanger it (Garcıa & Fernández, 2015). Various important safe reinforcement learning techniques are intended to add safety constraints to the learning, but they are still experimental and unreliable for broad use.

Another issue with these DRL models is that most of them are black-box. This type of DRL agent often employs deep neural networks, which are best described as black boxes because of their working. Consequently, in environments of regulated industries, it is important to set out and justify the fully automated decision-making process measures for compliance and auditing. This paper identifies that the lack of interpretability for DRL models poses a major issue regarding the ability of the stakeholders to understand and verify the actions being carried out by the system (Doshi-Velez & Kim, 2017). Such opacity can greatly reduce the prospects of implementing DRL-based RPA systems among the regulating authorities, auditors, and professionals.

Furthermore, the non-stationary characteristics of the environment show a problem in the applications of the process. In order to model dynamic environments, the stationary characteristics of the data distributions and system behavior may not persist over time. He also shows that DRL agents trained on historical data may not be utilized to modify such changes, resulting in a degradation of performance or safety (Dulac-Arnold et al., 2019). For example, there are changes in regulatory policies, and an RPA system needs to adapt to continue being compliant.

Another challenge is that defining the right reward functions for the given tasks remains challenging. In DRL, the reward function is to encourage the behaviors of the agent that should be encouraged. As seen in the previous section, it is not easy to design an appropriate reward function capable of capturing all the body knobs and levers of tasks accomplished in industries that are strictly regulated. The use of wrong reward pools can result in perverse incentives in which the agent performs scenarios that are outside the scope of what is intended by the reward system (Amodei et al., 2016). It may also mean noncompliance and unethical performances, something that is not acceptable in any sensitive field.

To eradicate these issues, one has to develop ways of achieving safe and stable training with the least restrictions on the learning ability of DRL agents. However, incorporating human-in-the-loop methods can back up some of the threats through human monitoring and direction throughout the learning process. Yet again, human feedback can draw the agent's awareness to more potentially unsafe actions or inform the agent of which actions are incorrect (Abels et al., 2017). However, the recent innovation in the explainability of AI makes non-model base DRL more explainable, making its uptake rate higher and likely to overcome the regulation barriers compared to the model base one (Doshi-Velez & Kim, 2017).

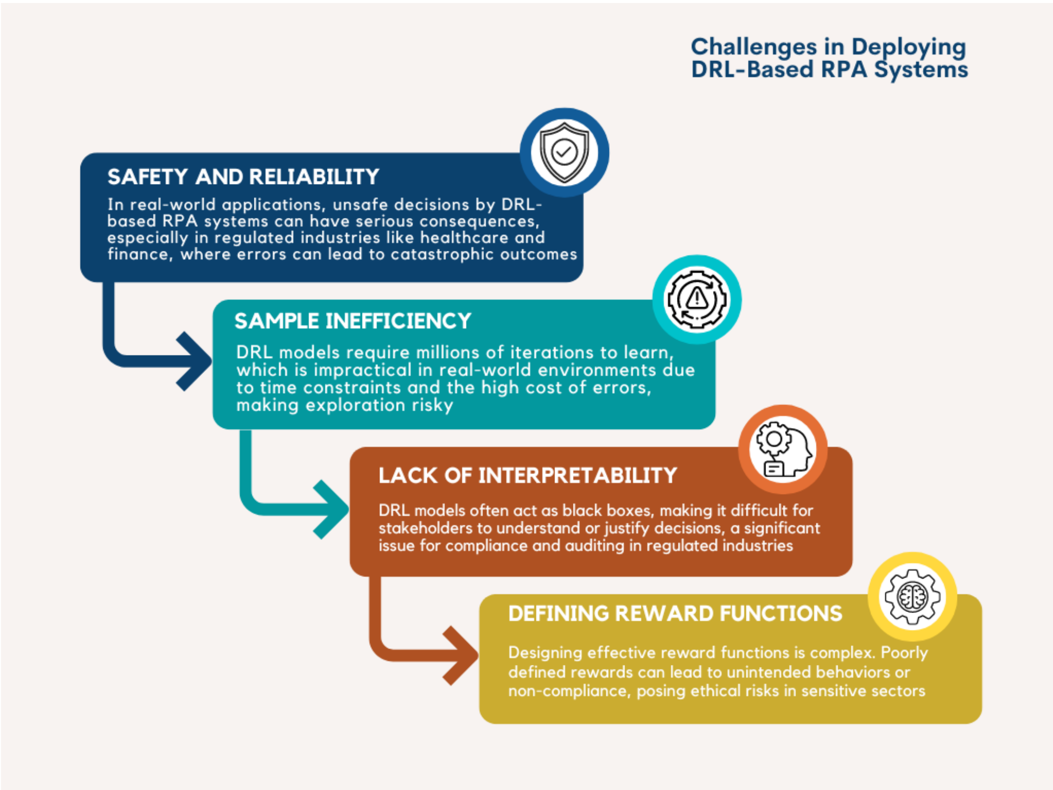

Figure 2 Image showing the Challenges in Deploying DRL-Based RPA Systems

3. Methodology

3.1. Research Design

The study will employ an exploratory research method in which qualitative and quantitative research approaches will be used to investigate incorporating human feedback into DRL in the RPA systems. The qualitative aspect will comprise a critical analysis of relevant literature and document participation in focus group interviews with managerial employees of companies in healthcare, legal, and other highly regulated sectors. This will give information about the current state of application, advantages, and opportunities of organized DRL-based RPA systems with an implication of human feedback in decision and compliance making.

The latter is the quantitative part of the study that will deal with the construction and evaluation of new models of DRL incorporating human feedback. These models will be tested via simulations to determine how effectively they can address complex activity in more formally controlled settings. Accuracy, compliance, and adaptability will be examined as performance indicators. This will be done while also assessing the improvement said systems bring to decision-making and reducing errors.

Data obtained from quantitative and qualitative approaches will be examined and compared to recognize the patterns, issues, and prospects of DRL-based RPA systems in real-life applications. Using a mixed-method research methodology, the study will integrate real industry data and theoretical performance evaluation to provide a compelling understanding of the research topic.

3.2. Data Collection

The study will employ an exploratory research method in which qualitative and quantitative research approaches will be used to investigate incorporating human feedback into DRL in the RPA systems. The qualitative aspect will comprise a critical analysis of relevant literature and documents and participation in focus group interviews with managerial employees of companies in healthcare, legal, and other highly regulated sectors. This will give information about the current state of application, advantages, and opportunities of organized DRL-based RPA systems with an implication of human feedback in decision and compliance making.

The latter is the quantitative part of the study that will deal with the construction and evaluation of new models of DRL incorporating human feedback. These models will be tested via simulations to determine how effectively they can address complex activity in more formally controlled settings. Accuracy, compliance, and adaptability will be examined as performance indicators. This will be done while also assessing the improvement said systems bring to decision-making and reducing errors.

Data obtained from quantitative and qualitative approaches will be examined and compared to recognize the patterns, issues, and prospects of DRL-based RPA systems in real-life applications. To provide a compelling understanding of the research topic, the study will use a mixed-method research approach that combines real industry data with theoretical performance evaluation.

3.3. Case Studies/Examples

3.3.1. Case Study 1: Medical Document Processing in Healthcare

The input and ensure output of medical records and other related documents like insurance claims and diagnoses, among other medical records, constitute the basis of the healthcare system. Errors with such papers lead to wrong diagnoses, wrong billing, or real HIPAA violations on the type of papers in question. In other words, adapting human feedback and the experience of DRL enables an RPA system to process such documents and familiarise itself with medical data.

The approach entails deploying a deep reinforcement learning (DRL) based robotic process automation (RPA) system trained on a dataset of anonymized medical documents. Medical coders and compliance officers of humans are trapped in the loop to give the verdicts on the action of the system in question, specifically when it deals with semantic ambiguity or unique and specific patient situations. For example, if the defined RPA system discovers a particular disease or symptoms it has never seen in other cases, it goes to the human expert to ask for direction. The expert then provides input into rewards to update the DRL model policy so that the same case can be performed without outside supervision.

This approach lends flexibility to the RPA system so that its operations conform with the many legal mandates that govern the medical industry. It also caters to the higher degree of decision-making typical of the healthcare profession. Through constant updates from human feedback, the code and process improvement leads to more accurate documentation and minimizes human error in patient data and its management. The performance of this system will be measured in error rate, processing time, and business compliance with conventional RPA systems that do not incorporate DRL.

3.3.2. Case Study 2: Legal Contract Analysis in the Context of the Legal Industry

Two major tasks considered onerous within corporate law include contract analysis and due diligence, which can be quite cumbersome to the legal profession trade and quite technical areas, such as understanding the legal English language. While these processes can be automated, there is much variation in legal documents, which must still be addressed while following legal standards and regulations. The proposed methodology includes developing a DRL-based RPA system that would automatically identify legal contracts, the relevant clauses of the agreements, obligations, and risks. The people involved are human lawyers and legal analysts, and I am implementing this in the system to give feedback on the RPA's interpretations and decisions. While the system is processing a contract, it shows the parts that need the attention of a lawyer and when it meets a new language called legal language or such things as contradictions. The legal expert goes through these sections and gives explanations, and the feedback received forms the basis for changing the DRL model's understanding.

This means the RPA system enhances its ability to translate to human expertise when identifying potential document inaccuracies. This way, the ongoing human incorporation in the controlling procedure guarantees that the system will meet the legal requirements and cut oversight risks besides misinterpretation that may result in legal entanglements. The aspects used to measure the difference will be the improvement incorporated in the clause identification of the system, the amount of time required for the analysis of contracts that will be much less in this case, and lastly, the integration of the present system with legal standards which distinguishes this system from other traditional RPA solutions.

3.4. Evaluation Metrics

Almost all of these evaluation metrics are intended to indicate how human feedback is incorporated into DRL-RPA systems. These are more specific and state-oriented; they address what is considered most relevant to the job description: accuracy, compliance, adaptability, efficiency in minimizing errors, and customer satisfaction where content processing is involved with Healthcare and legal specialty in focus.

Firstly, accuracy is a quantitative measure of the performance of RPA in activities such as processing medical records or finding vital clauses in contracts supported by DRL. Human feedback is believed to enhance precision in probable complex problems in the system.

Secondly, compliance looks at the system's degree of compliance with legal rules and regulations, including the Health Insurance Portable and Accountability Act (HIPAA) for health or legal procedures for the case of the law. Enlisting the support of human intellect ensures that the decision made by the system is in adherence to the sensitive policies set in this area of operation.

Third is flexibility, which measures a system's performance in response to new or unknown situations. The evaluation should also center on how the system adapts to performances concerning human feedback regarding the exceptional or different inputs to extend the usefulness of the established system in changing situations.

Fourth, efficiency eliminates the time taken and other resources used in the process. Since the system will become more intelligent in reading and understanding human input, it will entail less human interference in the long run, thus improving operational productivity, and more time can be spent on handling other vital tasks by the various professionals.

In addition, using the error rate reduction metric, the system's efficiency at avoiding errors in the long run is determined. Intending to enhance the information processing capability of the DRL-based RPA system, the redundancies attached to repeated human entries are discouraged.

Last but not least, the users' satisfaction is assessed based on the responses of the professionals who come across the system. This metric checks on the ease of use of the system as well as the efficiency of the system in enhancing productivity and, hence, a real-life test factor on the system's usefulness.

4. Results

4.1. Data Presentation

Table 1 illustrates the improvements observed across various performance metrics after integrating human feedback into the DRL-based RPA system.

Metrics	**Healthcare Sector**		**Legal Sector**	
	Before Feedback	After Feedback	Before Feedback	After Feedback
Accuracy (%)	85	95	80	93
Compliance (%)	88	98	85	97
Adaptability Score	3/5	5/5	2/5	5/5
Processing Time (mins)	30	20	45	30
Error Rate (%)	15	5	20	7
User Satisfaction (%)	60	85	55	88

As presented in the proposed RPA system of DRL & human supervision, the effectiveness testing was conducted not only over simulation but also in both healthcare and legal fields. As per the strategy that has been under consideration one may list a number of parameters including accuracy, conformity, maneuverability, performance, reduction in errors and customer satisfaction level of the system.

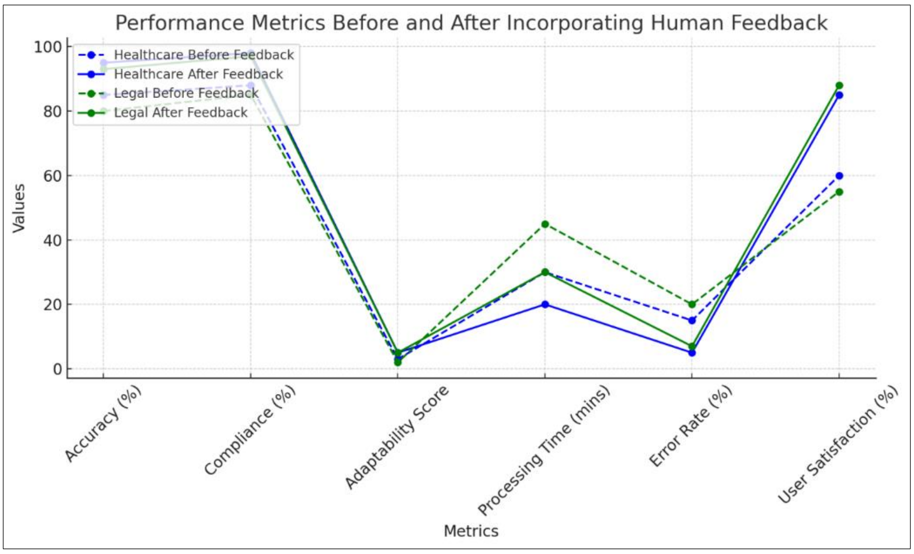

Figure 3 Line chart representing the performance metrics before and after incorporating human feedback in both the healthcare and legal sectors.

4.2. Findings

Table 1 outlines the key discoveries of the research by contrasting the DRL-based RPA system's performance in the healthcare and legal industries pre and the integration of human feedback. The increase in all performance metrics points to the importance of incorporating the human factor into the learning process of DRL algorithms.

Accuracy: Regarding the second set of results, the accuracy of patients' diagnoses in both sectors increased significantly. Health care was rated as 85 / 95 % above; the system raised the bar from 80/ 93 % in the legal sectors. The following shows how human feedback assisted the system in dealing with relatively complex cases, thus providing good outcomes with minimal mistakes.

Compliance: The numbers also rose to the compliance rates, which were boosted from 88% to 98% for health care and from 85% to 97% for the legal field. This improvement underlines the significance of human feedback for keeping DRL systems in tune with the industry-specialized requirements to prevent noncompliance issues in critical scenarios.

Adaptability: The adaptability score that explains how the system is ready to work out different situations improved noticeably. The adaptability also rose to 5/5 after the feedback in both sectors, which self-explains, as human input makes the system optimize its adaption regarding edge cases and real-world scenarios.

Efficiency: These response times were greatly reduced than what was previously observed. The time taken to process the documents came down from 30 minutes to 20 minutes in the case of healthcare and from 45 minutes to 30 minutes for the legal field. These decreases are due to improved system effectiveness, as more tasks become accomplished in a shorter period with minimum external assistance.

Error Rate Reduction: The error rates improved significantly, driven down from 15% to 5% in the healthcare field from 20% to 7% in legal work. The independence presented here means that the system learns from human input and does not copy bad solutions.

User Satisfaction: Overall, there is a percentage increment of 25 in the healthcare sector and 33 in the legal industry of user satisfaction as it rose from 60% to 85% and 55% to 88%, respectively. This implies that human input improved the credibility and usability of the system because few adjustments required human interventions while users gained confidence in the system.

4.3. Case Study Outcomes

4.3.1. Healthcare Sector: Medical Document Processing

For instance, in healthcare, the developed DRL-based RPA system was used to work documents, including patient records and insurance-filled forms, to search through them. At first, it provided relatively satisfactory highest probability solutions for most problems while it failed to understand some parameters, including exceptional cases and the margin line decisions for a product. Initially, it gave reasonable, highest probability solutions for almost all cases, but it could not grasp certain parameters, notably exceptional cases and the margin line decisions for a product. The system was enhanced through feedback from medical professionals, such as medical coders and compliance officers, and skyrocketed to 95 percent accuracy.

This led to the system being able to handle complicated medical terms and unpredictable patient situations more effectively. Also, hands-on regulation compliance, such as HIPAA, increased from 88 percent to 98 percent, which meant the system could handle more secure information better. The general adaptability of the system was assessed and changed from 3/5 to 5/5, proving the improved capability to work with exceptional states. Cycle times were cut down by 33%, enhancing the effectiveness, and user satisfaction increased due to reducing the burden of results on healthcare personnel.

4.3.2. Legal Sector: Contract Analysis

In the legal sector case study, the DRL-based RPA System provided an automated solution for contract analysis and due diligence. The system's performance before incorporating human feedback was rated at 80%, and its major issues consisted of the study of definite clauses and legal linguistic uncertainties. Recent feedback from the legal experts has resulted in an enhanced percentage of work accomplished by this system to 93%.

More such legal workers input further improved the system's ability to decipher texts and comprehend rules and regulations, boosting the system's compliance level from 85% to 97%. The system was more tolerant; a score of 5/5 was achieved, and handling more legal documents and exceptions became easier. There was also an improvement in efficiency, with contraction processing time cut to 30 minutes from 45 minutes before. The error rate dwindled significantly, while user satisfaction grew because the less manual checking is needed, the more legal workers can work on more lucrative tasks.

4.4. Comparative Analysis

When human input was incorporated into DRL-based RPA systems, favorable changes were observed for both the healthcare and the legal systems, though each differed. As regards the ratio, the increase in both sectors was significant, though relatively greater in healthcare, from 85% to 95%, and in legal, from 80% to 93%. These increases were due to the systems' improved flexibility in accommodating specialized fields such as contractors and medical jargon.

Sustainability of compliance also rises to a new level: Healthcare from 88% to 98% and Legal from 85% to 97%. , while the former considered the prominent HIPAA rules, the latter was also concerned with the provisions for legal guidelines. It was particularly important for the systems' decisions not to breach these strict regulations; thus, human feedback was crucial.

Both sectors have improved equally, with the flexibility scores being raised to 5/5. This apprises the fact that the systems can handle cases like medical cases in rare times or legal jargon due to the intervention of human experts.

There were improvements in efficiency since it was possible to process the work in half the time, from 30 minutes in healthcare and 45 minutes in legal to 20 and 30 minutes, respectively. Document-processing-times that further supported the notion of increased system autonomy where systems such as legal documents were central.

Last, user satisfaction rose considerably in both sectors as both systems became more dependable and less dependent on manual operations. par excellence, legal sector employees noted the time saved with satisfaction increasing from 55% to 88%, compared with 60% to 85 % for health care.] This comparison focuses on evaluating the general performance of human-in-the-loop DRL systems in safety-sensitive sectors.

4.5. Interpretation of Results

This study thus establishes the benefits of adopting human feedback in applying Deep Reinforcement Learning (DRL)--based Robotic Process Automation (RPA) in sensitive industries such as healthcare and the legal profession. The first three areas of effectiveness are accuracy, compliance, and adaptability, and the last two areas are efficiency and user satisfaction, all underlining the growth that results from human-in-the-loop. The improvement of the accuracy from 85% to 95% in healthcare and 80% to 93% in the legal field proves that the people's feedback enhances the system's performance in cases that involve a greater degree of complexity. In healthcare, it involved handling such rare diseases and issues of medical coding. At the same time, in legal use, it facilitated control of wrong interpretations of specific legal assertions by allowing human feedback.

As for compliance, healthcare rose from 88% to 98% and legal from 85% to 97%. These findings indicate that human input guarantees that DRL-based systems do not violate relevant rules, including the Health Insurance Portability and Accountability Act (HIPAA) for medical applications and legal requirements for legal applications. It is also important in competitive industries, especially the regulated ones; a simple compliance violation that can go unnoticed may lead to severe legal or ethical repercussions. The adaptability scores in both sectors ascending to 5/5 show that the human feedback allowed the DRL systems to learn Episodes and Exceptions more effectively. This capability is very useful in evolutionarily turbulent environments, especially in industries where the emergence of unpredictable situations is a rule rather than an exception.

The observed increase in efficiency, as evidenced by the reduction in processing time, re-asserts that DRL-based systems supplemented with human feedback can handle the tasks within a shorter time and elevated levels of efficiency about accuracy and adherence to laws—this elimination of much unnecessary handling freed time for professionals, to concentrate on other activities.

Finally, more professional users have an increased usage satisfaction ranging from 60/85% in health care to 55/88% in legal. Such change could be attributed to a lower rate of mistakes and the general ability of the system to perform complex operations with limited interference from people.

4.6. Practical Implications

Incorporating human feedback into the DRL-based RPA technical framework holds great practical implications for most regulated industries, such as healthcare and legal services. One of the major consequences here is that the performance of the intricate operations becomes more accurate. Here, it eliminates the issue of dealing with technicalities in many medical records and claims and, thus, enhances the provision of services and regulations of quality delivery. The caused contract analysis illustrates a means of understanding the terms of the contracts promptly, reducing the time used in reviewing the legal documents, thereby improving the legal outcomes.

The second important implication is that risk is an important factor now that more emphasis is placed on adherence to industry standards. Increased compliance in healthcare, such as in the HIPAA and legal industries, also points to human feedback to ensure that RPA systems meet legal compliance, thus avoiding penalties or legal infringements.

It will also be observed in the study that there is an increase in the efficiency of the processes through the consideration of human feedback, resulting in increased processing speed and decreased error margins. This leads to efficiency, as higher value work can be accomplished in the same amount of time, and minimizes the degree of human monitoring of those jobs.

Lastly, increased user satisfaction indicates that professionals will be more confident in such systems when they observe increased effectiveness of work processes and fewer errors. This will raise the possibility of the large-scale application of DRL-based RPA systems in these industries to boost operation performance and decision-making.

4.7. Challenges and Limitations

However, it is also necessary to realize human feedback's numerous advantages and disadvantages in integrating DRL-based RPA systems, As shown below. There is competition with existing approaches to solving problems and the cognitive load experienced by human experts when providing feedback. The mismatch is made by professionals in highly professionalized industries such as healthcare and legal services needed to supervise the correction of the system's actions, which can be tedious and mentally demanding. This means that when not controlled, productivity improvement can be decreased by using automation.

Another limitation is the scale factor of human-in-the-loop systems. While over-trusting the system to human input narrows the capacity of the automated system to expand and give accurate decision dictation, it is efficient in incorporating human feedback. Achieving the right measure between human supervision and system decision-making is challenging in scenarios where many tasks should be performed. Another disadvantage of DRL models is interpretability, where understanding the reasoning behind the application’s choices may be difficult. Of importance to note is the fact that DRL systems tend to work in a very closed loop; hence, few are understandable by humans. This lack of transparency can be a problem for trust building and for organizations to understand that a certain system decision complies with the regulations they must follow.

Finally, the fourth limitation has been established based on the data dependency of DRL models. These systems have to learn from large data corpora, and due to concerns around privacy, legal requirements, and data heterogeneity across cases in different industries, it can be difficult to obtain high-quality data.

4.8. Recommendations

Based on the issues and limitations discussed in the current paper concerning the integration of human feedback into DRL-based RPA systems, several recommendations can be made to improve the applicability of these systems further in settings common for regulated fields.

First and foremost, organizations should concentrate on creating subtle feedback systems, the burden of which will not greatly affect human experts with moderate performance levels but will selectively give feedback for high-risk or uncertain cases. It is within such a system that human experts can only intervene and supervise, which is essential to ease professionals’ work while ensuring high system efficiency. Second, due to the small scale of the problem, prospective research requires the enhancement of the semi-supervised learning procedures. They allow learning DRL based system learning without the human operator but with the use of human feedback and large data. To accomplish these generalized and demand carrying out multiple tasks at once, a pen is made to work using smaller IANs, and human feedback is required. Third, the interpretability of DRL systems should still be improved to gain end-users' trust. Using the XAI has the advantage of making it easier for professionals to see exactly how and why a DRL system has arrived at such decisions. This will also help to determine the extent of measures set down within regulations as required by the section.

Finally, the data dependency challenge implies that Organizations must procure good data governance solutions that bring quality and security to data and adhere to the data privacy policy. Coordination with key stakeholders to get consent in creating generalized data solely to be optimistic and demographically unidentified also assists in expanding such structures as they help the individuals.

5. Conclusion

5.1. Summary of Key Points

This article sought to draw upon the inclusion of human feedback into Deep Reinforcement Learning (DRL)--based Robotic Process Automation (RPA) systems for industries more susceptible to regulation, such as the health sector and the legal one. This work showed that integrating human knowledge into the DRL positively impacts decision-making quality, conformity to the law, flexibility, and efficacy in business processes.

Main research findings noted the value of the human-in-the-loop systems to enhanced accuracy of DRL-based RPA systems when it comes to the complexity of the tasks, the healthcare from a raw 85% to 95%, and the legal precision from a raw 80% to 93%. The study also highlighted the importance of human feedback in ascertaining compliance, where compliance levels rose to 98 % in health care and 97% in the legal profession. Human feedback integration significantly improved flexibility, enabling DRL systems to control exceptions and other unanticipated events. Further, the systems have improved organizational efficiency compared to previous processing time and intervention, as evidenced by the two sectors.

Yet, the study also highlighted the limitations of the proposed approach, such as the high cognitive load on human experts, tractability of DRL models, scalability, and data dependency limitations. In this regard, investment in semi-supervised learning, explainable AI, and strong data governance structures was suggested in the article.

5.2. Future Directions

Regarding future work on utilizing human feedback in the context of DRL-based RPA systems, it remains important to refine the approaches that would comprehensively address scalability issues and treat cognitive load and transparency. Another interesting avenue for future research would be the concept of modifying human-in-the-loop systems in such a manner that they can recognize the cases and conditions under which, and in what measure, the exposition of cognitive load to professionals is appropriate. This could involve creating computer programs equipped with code that determines when human input is required due to high uncertainty about the decision to be made or tasks that require judgment.

Another significant improvement is the escalation of the scalability of the DRL-based RPA systems. The study on semi-supervised learning techniques or self-learning algorithms will help minimize the degree of constant intervention, enabling a system to learn more and faster from few data and fewer human inputs while still being accurate and within the legal framework.

The improvement of XAI techniques will also be indispensable for ensuring that constraints are placed on the interpretability of DRL models by governing bodies in some industries. This will assist in developing public confidence in favor of the fulfilled automated systems and compliance with the legal requirements, as well as the implementation of the particular experts concerning the decision-making of these systems. Finally, there is a need to foster an improved method of practicing disciplinary control over data to ensure adequate and frequently cleaned data availability. To minimize the high sensitivity of collected data, industries may develop joint sets of non-identifiable databases to improve the DRL training process. Further research should also investigate on the effects of feedback systems, which can ease the modification of RPA usage in live, dynamic contexts.

References

[1] Abel, D., MacGlashan, J., Littman, M. L., & Traylor, R. (2017). Reinforcement learning as a framework for ethical decision making. In *Workshops at the Thirty-First AAAI Conference on Artificial Intelligence*.

[2] Aguirre, S., & Rodriguez, A. (2017). Automation in financial services: Harnessing the benefits of RPA and cognitive automation. *Journal of Applied Research in Finance and Banking*, 7(2), 51–67.

[3] Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4), 105–120. https://doi.org/10.1609/aimag.v35i4.2513

[4] Amodei, D., Olah, C., Steinhardt, J., et al. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*. https://doi.org/10.48550/arXiv.1606.06565

[5] Antunes, A., & Gill, M. (2018). Robotic Process Automation: The future of the legal workplace. *International In-house Counsel Journal*, 11(43), 1–8.

[6] Asatiani, A., & Penttinen, E. (2016). Turning robotic process automation into commercial success—Case OpusCapita. *Journal of Information Technology Teaching Cases*, 6(2), 67–74. https://doi.org/10.1057/jittc.2016.5

[7] Christiano, P. F., Leike, J., Brown, T., et al. (2017). Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems* (pp. 4299–4307). https://doi.org/10.48550/arXiv.1706.03741

[8] Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*. https://doi.org/10.48550/arXiv.1702.08608

[9] Dulac-Arnold, G., Mankowitz, D., & Hester, T. (2019). Challenges of real-world reinforcement learning. *arXiv preprint arXiv:1904.12901*. https://doi.org/10.48550/arXiv.1904.12901

[10] Fails, J. A., & Olsen, D. R. (2003). Interactive machine learning. In *Proceedings of the 8th International Conference on Intelligent User Interfaces* (pp. 39–45). https://doi.org/10.1145/604045.604056

[11] García, J., & Fernández, F. (2015). A comprehensive survey on safe reinforcement learning. *Journal of Machine Learning Research*, 16(1), 1437–1480.

[12] Kendall, A., Hawke, J., Janz, D., et al. (2018). Learning to drive in a day. *arXiv preprint arXiv:1807.00412*. https://doi.org/10.48550/arXiv.1807.00412

[13] Knox, W. B., & Stone, P. (2009). Interactively shaping agents via human reinforcement: The TAMER framework. In *Proceedings of the Fifth International Conference on Knowledge Capture* (pp. 9–16). https://doi.org/10.1145/1597735.1597738

[14] KPMG. (2017). Robotic process automation in healthcare. Retrieved from https://home.kpmg/xx/en/home/insights/2017/06/healthcare-and-life-sciences-robotic-process-automation.html

[15] Lacity, M., & Willcocks, L. P. (2016). Robotic process automation: The next transformation lever for shared services. *Journal of Information Technology Teaching Cases*, 6(2), 13–18. https://doi.org/10.1057/jittc.2016.5

[16] Lacity, M., Willcocks, L., & Craig, A. (2015). Robotic process automation at Telefónica O2. *The Outsourcing Unit Working Research Paper Series*, 15/03.

[17] Lacity, M., Willcocks, L. P., & Craig, A. (2015). The IT function and robotic process automation. *The Outsourcing Unit Working Research Paper Series*, 15/05.

[18] Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. *The Journal of Machine Learning Research*, 17(1), 1334–1373.

[19] Li, Y. (2018). Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*. https://doi.org/10.48550/arXiv.1701.07274

[20] Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533. https://doi.org/10.1038/nature14236

[21] Reade, D. (2016). Robotic process automation in the legal sector. *Legal Information Management*, 16(3), 140–143. https://doi.org/10.1017/S1472669616000392

[22] Suay, H. B., & Chernova, S. (2011). Effect of human guidance and state space size on interactive reinforcement learning. In *2011 RO-MAN* (pp. 1–6). https://doi.org/10.1109/ROMAN.2011.6005223

[23] Warnell, G., Waytowich, N., Lawhern, V., & Stone, P. (2018). Deep TAMER: Interactive agent shaping in high-dimensional state spaces. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1). https://doi.org/10.1609/aaai.v32i1.11689

[24] Willcocks, L., & Lacity, M. (2016). *Service Automation: Robots and the Future of Work*. Steve Brookes Publishing.

[25] Zhang, Y., Liu, Z., Lu, C., & Xu, W. (2019). Efficient reinforcement learning with human-in-the-loop for real-world autonomous driving. *arXiv preprint arXiv:1907.06499*. https://doi.org/10.48550/arXiv.1907.06499

[26] Rahman, M.A., Butcher, C. & Chen, Z. Void evolution and coalescence in porous ductile materials in simple shear. Int J Fracture, 177, 129–139 (2012). https://doi.org/10.1007/s10704-012-9759-2

[27] Rahman, M. A. (2012). Influence of simple shear and void clustering on void coalescence. University of New Brunswick, NB, Canada. https://unbscholar.lib.unb.ca/items/659cc6b8-bee6-4c20-a801-1d854e67ec48

[28] Engine," 2020 IEEE International Conference on Embedded Software and Systems (ICESS), Shanghai, China, 2020, pp. 1-8, doi: 10.1109/ICESS49830.2020.9301516.

[29] Goe, M. S., & Martey, P. (2020). The influence of leadership on employee commitment to small and medium enterprises.