

Distributive Training Can Improve Neural Network Performance based on RL-CNN Architecture

Dmytro Proskurin¹, Sergiy Gnatyuk¹, and Madina Bauyrzhan²

¹National Aviation University, 1 Liubomyr Huzar ave., 03058, Kyiv, Ukraine

²Satbayev University, 22 Satbayev str., Almaty, Kazakhstan

Abstract

Recent studies of reinforced learning have suggested a distributive method of rewarding, which can be used to improve the performance of the deep Q-learning networks for such tasks as image classification, item recognition and movement simulation. This article uses the same method for more basic neural network types (RL + CNN more specifically) to conclude, whether this approach in machine learning systems can improve the results without the need for profound architectural changes or test and dev data cleansing. The same technique can then be used to improve the AI-SIEM threat detection system.

Keywords

Neural networks, reinforcement learning, deep Q-learning, dopamine, convolutional neural network, recognition agents

1. Introduction

Today, neural networks are used to solve many business problems, such as sales forecasting, customer research, data validation, risk management, anomaly detection, and even natural language comprehension. They may be the part of technological development that currently has the greatest potential. In addition to the generic architecture, there are integrated implementations using RL, or enhanced training, that address which actions, the intelligent agents should perform in a particular environment to maximize the notion of aggregated reward.

In this paper we describe how updating the reinforcement learning (RL) algorithm in the model with Distributional RL improves the performance without any deeper changes in the NN architectures. We use and update an effective aircraft detection framework based on reinforcement learning and convolutional neural network (RL-CNN) from [18]. The process of localization aircraft can be seen as an action decision problem with a sequence of actions to refine the size and position of bounding box. Active interaction to understand the image region, change of the correct bounding box aspect ratio and selection region of interest are important to determine the accurate position of aircraft. Based on the characteristics of reinforcement learning and our specific aircraft localization process, we use reinforcement learning to learn and implement our framework. Compared with the object detection method based on reinforcement learning, our aircraft detection framework combines the advantages of reinforcement learning and supervised learning, and is able to detect unfixed number of aircraft in remote sensing images. In addition, compared with the structured prediction bounding box regression algorithms, our detection agent dynamically localizes aircraft [18].

Our work has the following contributions:

1. We combine the distributional reinforcement learning and supervised learning in our aircraft detection framework. We train the aircraft detection agent with the deep Q learning method and

CPITS-II-2021: Cybersecurity Providing in Information and Telecommunication Systems, October 26, 2021, Kyiv, Ukraine
EMAIL: proskurin.d@stud.nau.edu.ua (D. Proskurin); s.gnatyuk@nau.edu.ua (S. Gnatyuk); madina890218@gmail.com (M. Bauyrzhan)
ORCID: 0000-0003-1336-6055 (D. Proskurin); 0000-0003-4992-0564 (S. Gnatyuk); 0000-0002-8287-4283 (M. Bauyrzhan)



© 2022 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

train the CNN model with supervised learning to learn the appearance characteristics of the aircraft.

2. We train the detection agent with distributional reinforcement learning and apprenticeship learning, which guide the detection agent with the greed strategy.
3. The proposed approach improves the running time of the aircraft detection framework without any deep changes.
4. The mentioned changes can be then applied to the AI-SIEM system based on a combination of event profiling to improve the overall threat recognition and performance rates.

2. Related Work

The dopamine-based reinforcement learning architecture is used in the deep Q-learning neural network in the “A distributional code for value in dopamine-based reinforcement learning” article by W. Dabney [3]. The article shows the advantages of such architecture (Figure 1) for the tasks as image recognition and character movement simulation proving it to be better than the typical deep Q-learning networks. But the article does not provide any conclusion on whether the dopamine-based learning can be used in any primitive (basic) neural network types or for any other tasks. This is the reason, why we decided to try and use this architecture in basic architectures (e.g., CNN) to compare its results for the basic tasks.

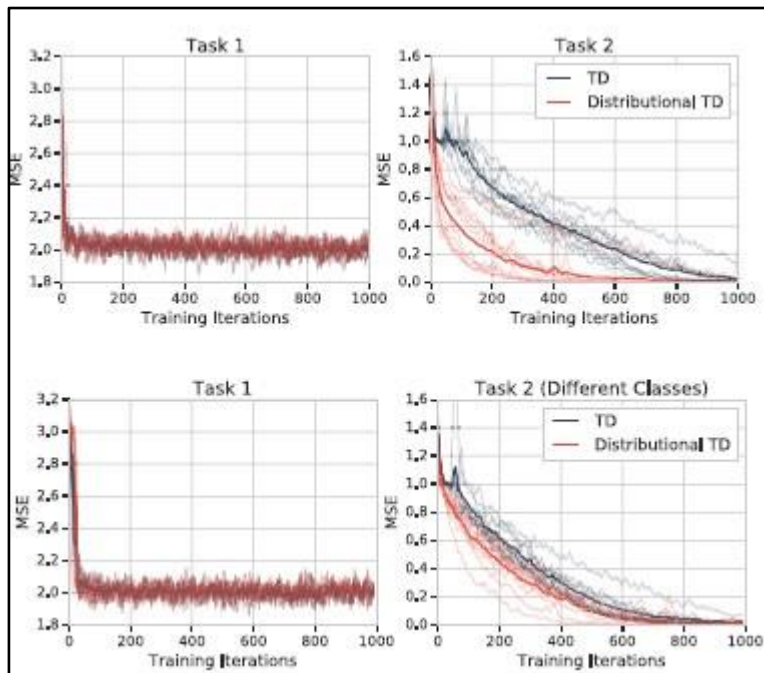


Figure 1: Simulation experiment to examine the role of representation learning in distributional RL

Long Wen et al. [16] adopted a new learning rate scheduler based on the reinforcement learning (RL) for convolutional neural network (RL-CNN) in fault classification, which can schedule the learning rate efficiently and automatically. The new RL agent is designed to learn the policies about the learning rate adjustment during the training process, changing the RL-CNN structure and testing it afterwards. This approach can be used to minimize the faulty classification in the aircraft detection model described in this paper.

Yang Li et al. [18] proposed an effective aircraft detection system based on reinforcement training and a convolutional neural network (CNN) model. The aircraft in the images can be accurately and reliably located using a search engine, so that the place-candidate is dynamically reduced to the correct location of the aircraft, which is realized by training via reinforcements. The detection framework overcomes the difficulty that modern detection methods based on reinforcement training can detect only

a fixed number of objects. They use a limited EdgeBox, which first creates quality candidate boxes with prior knowledge of aircraft. They then train the intelligent detection agent through advanced training and learning. The detection agent accurately identifies the aircraft in the candidate boxes in a few steps, and it even works better than the greed strategy in the learning process. At the final stage of detection, we carefully develop the CNN model, which assumes that the localization result obtained by the detection agent is an aircraft. Comparative experiments demonstrate the accuracy and efficiency of our aircraft detection system [20].

Leef Jonghoon et al. [19] developed an AI-SIEM system based on a combination of event profiling for data preprocessing and different artificial neural network methods, including FCNN, CNN, and LSTM. The system focuses on discriminating between true positive and false positive alerts, thus helping security analysts to rapidly respond to cyber threats. All experiments in this study are performed by authors using two benchmark datasets (NSLKDD and CICIDS2017) and two datasets collected in the real world.

3. Description of the Distributive Approach

In this section, we offer a more complete and clearer introduction to distributive learning. Since its introduction, the theory of dopamine reward prediction errors has explained a large number of empirical phenomena, providing a unifying basis for understanding the representation of reward and value in the brain. According to current canonical theory, reward predictions are presented as a single scalar quantity that supports the knowledge of expectations or the mean value of stochastic results. Here, we offer a description of dopamine-based reinforcement training inspired by recent artificial intelligence research on distributive reinforcement training. We assume that the brain represents possible future rewards not as a single mean, but as a distribution of probabilities, effectively representing many future outcomes simultaneously and in parallel. This idea includes a set of empirical predictions that we tested using single records of the ventral tegmental region of the mouse. Our findings provide strong evidence for the neural implementation of distribution reinforcement training.

The problem of RL is to find the best way of learning, which is the characteristic of each stage, and what actions should be taken at this stage. Thus, starting from state x , the start of training determines the future rewards (R_t) that will be received for all future processes (t). The sum of these future rewards is called income, Z (1):

$$Z(x) = \sum_{t=0}^{\infty} \gamma^t R_t \quad (1)$$

However, Z is not a single number, but a random variable, the value of which depends on possible rewards. Given the known initial state (x) and behavior, there are two sources of this coincidence: the environment may have stochastic transitions and rewards, and the behavior itself may be probabilistic. Most existing RL algorithms learn about $V = E[Z]$, which is an expectation of the result. For example, a simple rule TD with learning speed (α):

$$\delta = r + V(x') - V(x), V(x) \leftarrow V(x) + \alpha \delta \quad (2)$$

δ is calculated during the transition from state x to state x' , moves V in the direction of waiting for the result. However, it may be useful to know the full distribution of these results. There is an easy way to study this distribution using an exceedingly small modification of the standard TD. In this method, instead of one value function, a complete set of value functions is studied. For each function of the value V_i , a separate reward prediction error is calculated, where $V_j(x')$ is a sample from the distribution $V(x')$. The update of the values of the first equation is modified so that different learning indicators (positive and negative) are applied to positive and negative RPE:

$$V(x) \leftarrow V(x) + \alpha_i \delta_i \quad \text{for } \delta_i > 0 \quad (3)$$

$$V(x) \leftarrow V(x) + \alpha_i^- \delta_i \quad \text{for } \delta_i < 0 \quad (4)$$

The studied V_i together make a set of sufficient data for the distribution of results from the state x . Even though population data on average mimic the classical function of the mean, they differ significantly. Thus, distributed training arises automatically from a quite simple modification to standard TD training. The learning rule described in equations 2 and 3 can be summarized as follows:

$$V(x) \leftarrow V(x) + \alpha_i^+ f(\delta_i) \quad \text{for } \delta_i > 0 \quad (5)$$

$$V(x) \leftarrow V(x) + \alpha_i^- f(\delta_i) \quad \text{for } \delta_i < 0 \quad (6)$$

where f is the function that converts the prediction error. For any non-decreasing f , the studied V_i compile a set of statistics for the distribution of results. This is especially easy to see in the special case $f(x) = \text{sign}(x)$. Let $\alpha_i^+ = 3$ and $\alpha_i^- = 1$. In other words, each positive prediction error (PE) is +3 and each negative PE is -1. In this case, V_i is calculated to accurately predict the upper quartile of the result distribution. In the upper quartile of negative PE will be three times more than positive PE, so after weighing on α_i^+ and α_i^- positive and negative PE will be exactly balanced.

Thus, this is the main point of learning dynamics. Under the basic assumptions, each V_i converges exactly on the quantile of the result distribution. A specific quantile is.

$$\alpha_i^+ T_i = \alpha_i^- (1 - T_i) \quad (7)$$

In other words, V_i is a result that has a cumulative probability i . While T_i covers the range from 0 to 1, the set V_i together forms a complete quantile function (also known as the inverse cumulative distribution function) for the distribution of data (positive and negative) [3].

3.1. Basic Project Architecture based on RL-CNN

Reinforced training is aimed at reward training and consistent decision-making. It is widely used in various applications such as robotics management, treatment, finance and games. For the first time, DeepMind proposed a combination of learning reinforcement and a deep neural network to play Atari 2600 video games. In some games, a q-learning agent achieves superhuman performance. In addition, AlphaGo won the Go competition, which has been studied by professional players for thousands of years. An agent who has been trained in training algorithms can directly control autonomous helicopters.

Recent research focuses on discrete or control, value function or policy. In contrast to controlled learning methods, reinforced learning approaches are attracting more and more attention. Reinforcement training proves a new method for solving traditional computer vision problems, such as visual tracking, action recognition and object detection, using deep reinforcement algorithms.

Over the last two years, some methods of object detection have been proposed based on reinforcement training. Having implemented such an algorithm, the agent is taught the policy of determining a fixed number of objects by gradually narrowing the limiting frame from the image of the hole to the main object using actions that can change the scale, location or have a restrictive framework. Some studies in [2] reduced the number of actions to six (Figure 2), and this facilitated the optimization policy. Similar to [1], the detection system in [2] also used the return inhibition mechanism to determine a fixed number of objects.

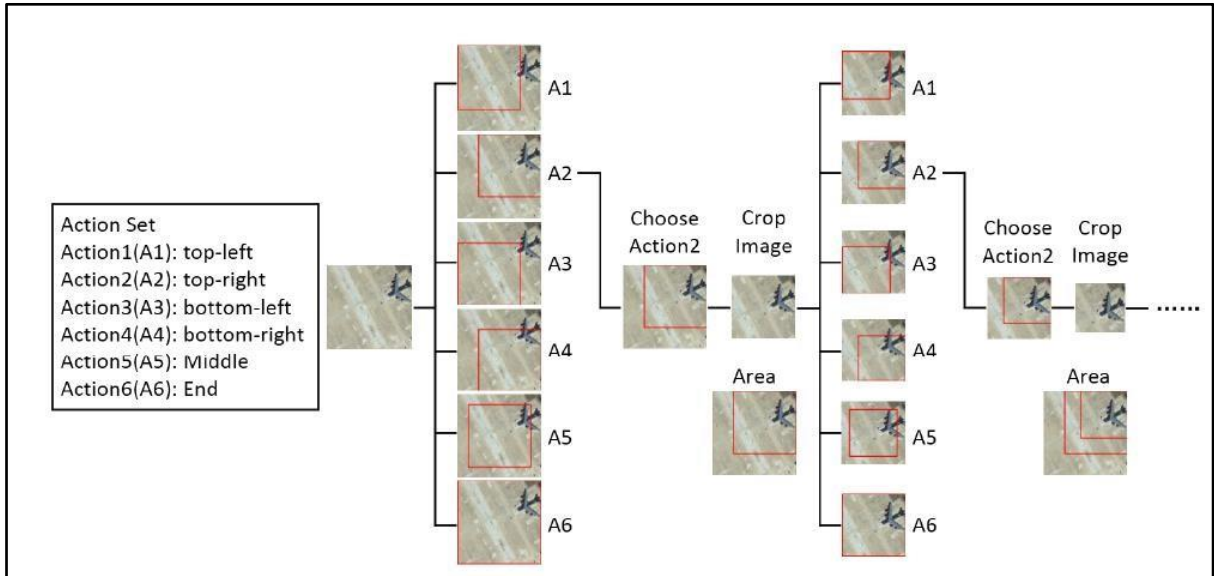


Figure 2: Search process of the detection agent

In contrast to [1], the agent [2] adopted a hierarchical representation, which performed a top-down search for objects. In [4], a new method based on reinforcement training for efficient generation of object propositions is proposed: agent balanced the localization of covered objects and the disclosure of uncovered ones with an effective reward function.

With the development of remote sensing technology and the improvement of resolution, automatic detection of aircraft not only plays an important role in the military industry, but also in the field of civil aviation. This process is one of the important areas of research in the analysis of images using remote sensing. Methods for detecting objects using convolutional neural networks (CNN) have recently been proposed. Using these state-of-the-art methods, the detection of aircraft in remote sensing images was developed: it became possible to derive the coordinates of aircraft through the regression of neural networks [5].

We force the detection agent to interact with 2700 images. When the detection agent finishes interacting with all 2700 remote sensing images, one training epoch ends and we train the agent for 30 epochs. At the learning stage, we use ϵ - greedy policy. With the beginning of training, the value of ϵ decreases from 0.9 to 0.1, which means that at the beginning of training, the detection agent has a high probability of 0.9 to random actions to study various conditions. The value of ϵ decreases by 0.1 when one epoch ends, and finally it is fixed to 0.1. In later epochs, the detection agent more often chooses actions based on trial and error to use the experience learned by him.

We teach two types of detection tools. The first agent chooses random actions at the intelligence stage, and we call this an agent without knowledge. On the other hand, we know the truth of the plane at the training stage, so the second agent chooses actions based on a strategy of greed in the training process, teaches the agent whose action is best to get the most insertion-over-union gain. We call this an agent with knowledge. Each type of detection agent performs a total of 120 million actions over 30 epochs, and the deep Q network is updated 120 million times. The detection agent performs an average of five actions to locate the aircraft.

Some papers perceived the problem of object detection as the Markov decision-making process and prepared an RL-based detection agent. The detection agent takes a top-down search process that first analyzes the global image and then gradually narrows the local regions that contain information about the object. However, these detection methods, based on reinforcement training, detect only a fixed number of objects, and cannot solve the problem of detecting aircraft in remote sensing images. This paper uses an aircraft detection system based on RL and the CNN model (RL-CNN), which is shown in Figure 3. The aircraft localization process can be considered as a decision-making problem with a sequence of actions to clarify the size and position of the boundary frame. The image area understanding

function must be able to change the boundaries of the frame and the selection area to determine the exact position of the aircraft. Based on the characteristics of reinforcement training and our specific aircraft localization process, we use RL to study and implement this structure. Compared to other methods of object detection, our aircraft detection system combines the advantages of RL and supervised learning and can detect an unspecified number of aircraft in remote sensing images [1, 5].

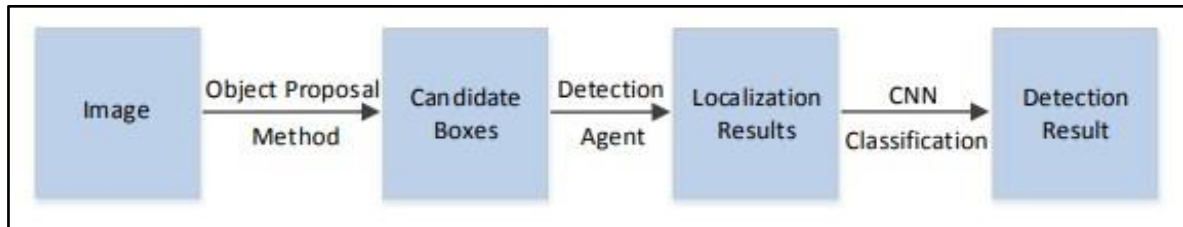


Figure 3: Reinforcement learning and CNN (RL-CNN) aircraft detection framework

In this detection structure, the object proposal method is used to create the frames of the candidate aircraft in the original image. Training the detection agent by RL places the aircraft within this framework. Then they narrow around the plane. The CNN model evaluates each localization result, and we can detect any number and type of aircraft in remote sensing images after maximum narrowing [6, 7].



Figure 4: Data example for recognition

This work has the following contributions:

- We combine RL and controlled learning in our system and train agents using Deep Q-learning and CNN to learn the characteristics of all aircraft in the image.
- We train RL-based agents who manage them with a very "greedy strategy". The detection agent even works more efficiently than a similar approach in some test samples.



Figure 5: Agent recognition of the individual structures

- This system overcomes the difficulty of having a modern RL-based detection method that can only detect a fixed number of objects (where classification is based on the basic images already studied). In remote sensing images, we can detect any number and types of aircraft by narrowing, classifying, and studying.

This architecture offers a new aircraft detection system based on RL-CNN using remote sensing. Limited EdgeBox generates high quality, but small number of candidates based on previous knowledge of aircraft. An intelligent detection agent based on RL learns by mastering a new state and using his own experience, and he can accurately find the plane in the fields step by step, following the search behavior from top to bottom. Using a combination of detector models and CNN models, which assume that the localization of the detection agent is likely to result in an aircraft, this RL-CNN system can detect the aircraft in remote sensing images (Figs. 4, 5).

3.2. Introduction of DRL Approach

The implementation of a distributive approach requires changes from the first part to the RL stage of the architecture. This implementation improves the efficiency of artificial agents by stimulating the learning of richer "ideas". The conditions associated with the same expected profit, but a wide distribution of results, can be represented as different under the DRL, but not under the usual RL.

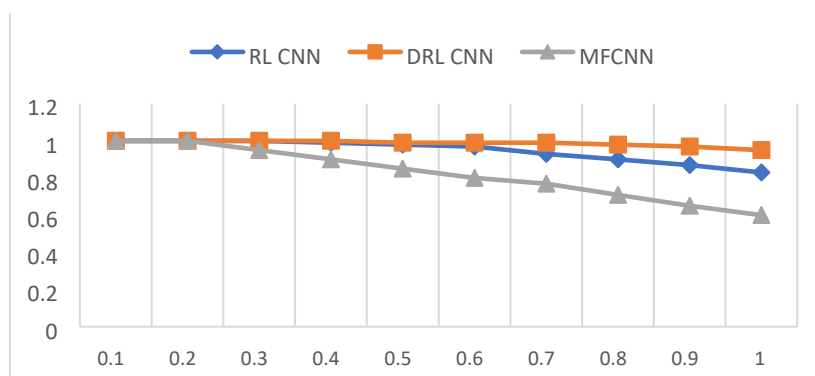


Figure 6: Results of three architectures using the same training data (Precision – Recall Curves)

We compared the DRL CNN model with the basic RL CNN and Multi-model Fast Regions CNN (MFCNN) ones.

Experiments show that the detection agent in our DRL-CNN can give up immediate interests and focus on long-term reward to get great results. DRL-CNN aircraft detection structures can not only

better detect an unfixed number of aircraft in remote sensing images, but also require less time and data to obtain a similar result (compared to RL CNN and MFCNN) based on the same test training set (Figure 6).

We compare the detection time on each testing image for four aircraft detection frameworks. Compared with MFCNN, our DRL-CNN generates less candidate boxes. Thus, DRL-CNN is faster than MFCNN and basic RL-CNN. Benefitting from the DRL, DRL-CNN can faster identify and process the input remote sensing image, and it needs less time than MFCNN and basic RL-CNN. Therefore, we can see that, the DRL-CNN requires less time to identify the object and has a smaller chance of faulty classification which leads to a more precise result.

As a result, the overall performance of the RL+CNN architecture is better on the same dataset and additional parameters. The result has been replicated several times to confirm the improvement.

The [3] suggests, that DRL improves performance because it drives learning of richer representations. States associated with the same expected return, but different return distributions, may be represented as different under distributional RL, but not under ordinary RL. (A similar proposal has been made about the performance benefits that arise from requiring an agent to learn multiple value functions at multiple discount rates).

To illustrate this idea, they analyzed the representations learned by DQN and distributional TD in the Atari 2600 game Ms. Pacman. For both fully trained agents they generated trajectories by allowing 2 the agents to play the game, but with a higher than usual (0:1) probability of taking random actions. For each trajectory they recorded the full-resolution frame and the activations of the final hidden layer of the neural network (a 512 dimension real-valued representation vector). They then trained, for each agent, a linear decoder from the representation vector to the game frame. How well the agent can reconstruct the full game state can be taken as an indication of how rich a representation has been learned. They split each agent's trajectories into a training and testing set.

As a result, the paper suggests that the DRL approach is an overall improvement and the case based one, which leads to a conclusion that any NN + RL architecture can be improved without requirement any deep changes.

4. Introduction of the DRL Model into the AI-SIEM based on the FCNN, CNN, and LSTM

Figure 7 shows the system architecture of the big data platform from Lee J. et al. [19]. The platform mainly consists of a data collection system, data processing system, data analysis and data storage system to analyze cyber-threat information using long-term security data. Using the techniques for large-scaled data processing, this platform is capable of continually collecting the numerous streamed security events and processing the data in real-time. Based on the big data platform, the proposed methods can be coupled with AI-based SIEM. By adopting AI technique to the platform, true alerts can be better differentiated from false alerts in the real world [19].

In this paper, they have proposed the AI-SIEM system using event profiles and artificial neural networks. The system enables the security analysts to deal with significant security alerts promptly and efficiently by comparing long-term security data. By reducing false positive alerts, it can also help the security analysts to rapidly respond to cyber threats dispersed across a large number of security events. For the evaluation of performance, we performed a performance comparison using two benchmark datasets (NSLKDD, CICIDS2017) and two datasets collected in the real world. First, based on the comparison experiment with other methods, using widely known benchmark datasets, they showed that our mechanisms can be applied as one of the learning-based models for network intrusion detection. Second, through the evaluation using two real datasets, they presented promising results that our technology also outperformed conventional machine learning methods in terms of accurate classifications. In the future, to address the evolving problem of cyber-attacks, we will focus on enhancing earlier threat predictions through the multiple deep learning approach to discovering the long-term patterns in history data. In addition, to improve the precision of labeled dataset for supervised-learning and construct good learning datasets, many SOC analysts will make efforts directly to record labels of raw security events one by one over several months [19].

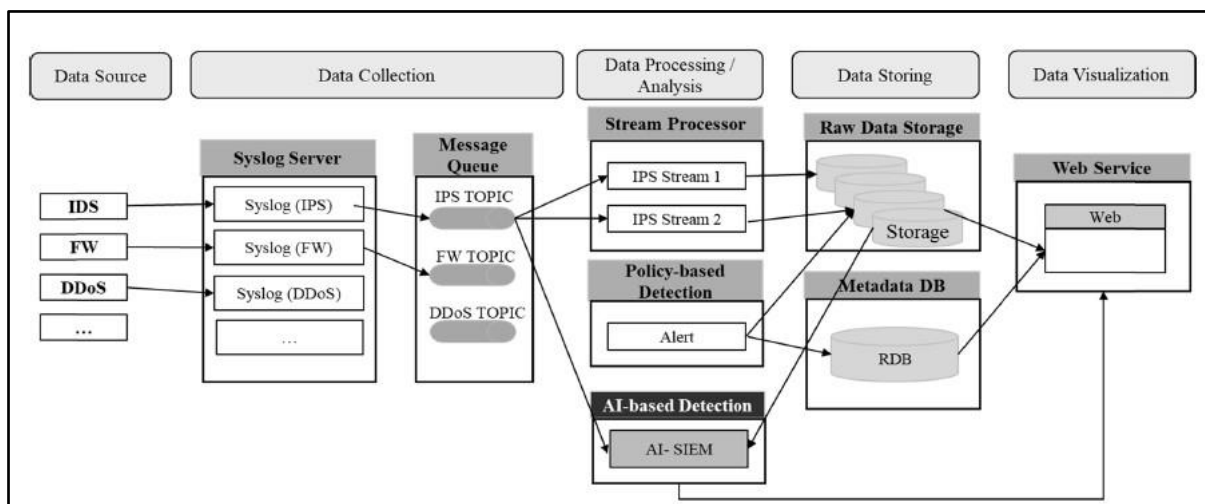


Figure 7: The architecture of our big data platform for AI-based SIEM

Therefore, the DRL can be used for prior classification of the IPS outputs creating a proceeded input for the neural networks. This approach can significantly increase the processing times and the overall results of the system. The results from this paper and related works showed, that the DRL algorithm improves both the data processing and training times for the CNN-based models, which will lead to a higher rates and less faulty classifications. But, further research is required.

5. Conclusion

In this paper, we use an effective and tested RL-CNN aircraft detection framework based on reinforcement learning and the CNN model. The EdgeBoxes are able to precisely identify the aircraft to pass them to the CNN. The detection agent based on reinforcement learns through exploration of new state and exploitation of its own experience, and it can accurately locate the aircraft in the candidate boxes step by step following the top-down searching policy. From the results we can see that the implementation of distributive learning formulas in standard RL-CNN architecture improves the basic characteristics of aircraft detection by the framework This update can be used for the RL + NN architecture, without the need for profound changes to the RL phase. Thus, the amount of required training data is reduced, which leads to faster learning and updating of basic parameters, which can affect the creation of more compact systems. As a result, the DRL algorithm can be implemented into the developed AI-SIEM system to classify the IPS outputs prior neural networks training iterations. Which leads to lower processing times and data requirement.

Despite the better performance there are some shortcomings in our work. Some additional research must be done to determine the long-term effects and whether the dataset has any effect on the outcome. We consider exploring how the DRL approach affects the model for other tasks such as language processing.

6. References

- [1] J. C. Caicedo, S. Lazebnik, Active object localization with deep reinforcement learning. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 11–18 December 2015; pp. 2488–2496.
- [2] Bellver, M.; Giró i Nieto, X.; Marqués, F.; Torres, J. Hierarchical Object Detection with Deep Reinforcement Learning. arXiv 2016, arXiv:1611.03718. Available online: <http://arxiv.org/abs/1611.03718> (accessed on 1 February 2018).

- [3] Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., & Botvinick, M. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature*, 577 (7792), 671-675.
- [4] Jie, Z.; Liang, X.; Feng, J.; Jin, X.; Lu, W.; Yan, S. Tree-structured reinforcement learning for sequential object localization. In *Proceedings of the Advances in Neural Information Processing Systems*, Barcelona, Spain, 5–10 December 2016; pp. 127–135.
- [5] Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In *European Conference on Computer Vision*; Springer: Amsterdam, The Netherlands, 2016; pp. 21–37.
- [6] Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. In *Proceedings of the Advances in Neural Information Processing Systems*, Montreal, QC, Canada, 7–12 December 2015; pp. 91–99.
- [7] Wu, H.; Zhang, H.; Zhang, J.; Xu, F. Fast aircraft detection in satellite images based on convolutional neural networks. In *Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP)*, Quebec City, QC, Canada, 27–30 September 2015; pp. 4210–4214.
- [8] Silver, D.; Lever, G.; Heess, N.; Degris, T.; Wierstra, D.; Riedmiller, M. Deterministic policy gradient algorithms. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, Beijing, China, 21–26 June 2014; pp. 387–395.
- [9] Yun, S.; Choi, J.; Yoo, Y.; Yun, K.; Choi, J.Y. Action-Decision Networks for Visual Tracking with Deep Reinforcement Learning. In *Proceedings of the 2017 Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21–26 July 2017; pp. 1349–1358.
- [10] Jayaraman, D.; Grauman, K. Look-ahead before you leap: End-to-end active recognition by forecasting the effect of motion. In *European Conference on Computer Vision*; Springer: Amsterdam, The Netherlands, 2016; pp. 489–505.
- [11] Hosang, J.; Benenson, R.; Dollár, P.; Schiele, B. What makes for effective detection proposals? *IEEE Trans. Pattern Anal. Mach. Intell.* 2016, 38, 814–830.
- [12] Cheng, M.M.; Zhang, Z.; Lin, W.Y.; Torr, P. BING: Binarized normed gradients for objectness estimation at 300fps. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 23–28 June 2014; pp. 3286–3293.
- [13] Uijlings, J.R.; Van De Sande, K.E.; Gevers, T.; Smeulders, A.W. Selective search for object recognition. *Int. J. Comput. Vis.* 2013, 104, 154–171.
- [14] Zitnick, C.L.; Dollár, P. Edge boxes: Locating object proposals from edges. In *European Conference on Computer Vision*; Springer: Zurich, Switzerland, 2014; pp. 391–405.
- [15] Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; et al. Human-level control through deep reinforcement learning. *Nature* 2015, 518, 529–533.
- [16] L. Wen, X. Li and L. Gao, "A New Reinforcement Learning Based Learning Rate Scheduler for Convolutional Neural Network in Fault Classification," in *IEEE Transactions on Industrial Electronics*, vol. 68, no. 12, pp. 12890-12900, Dec. 2021, doi: 10.1109/TIE.2020.3044808.
- [17] M. Karimzadeh, A. Esposito, Z. Zhao, T. Braun and S. Sargento, "RL-CNN: Reinforcement Learning-designed Convolutional Neural Network for Urban Traffic Flow Estimation," 2021 *International Wireless Communications and Mobile Computing (IWCMC)*, 2021, pp. 29-34, doi: 10.1109/IWCMC51323.2021.9498948.
- [18] Li, Y.; Fu, K.; Sun, H.; Sun, X. An Aircraft Detection Framework Based on Reinforcement Learning and Convolutional Neural Networks in Remote Sensing Images. *Remote Sens.* 2018, 10, 243. <https://doi.org/10.3390/rs10020243>.
- [19] Lee, J.; Kim, J.; Kim, I.; Han, K. Cyber Threat Detection Based on Artificial Neural Networks Using Event Profiles. *IEEE Access* 2019, 7, 165607–165626.
- [20] V. Buriachok, et al., Invasion Detection Model using Two-Stage Criterion of Detection of Network Anomalies, *Cybersecurity Providing in Information and Telecommunication Systems (CPITS)*, pp. 23–32, Jul. 2020.