Cognitive Radar: Literature Survey

Ajay Ramesh Ranganathan (IMT2017502)

NC863 Course Project

International Institute of Information Technology, Bangalore

May 22nd 2021

1 Overview

Modern radar systems face a variety of challenges due to the ever increasing requirements of robustness, high performance and flexibility in real world scenarios. A type of radar called the cognitive radar has shown promise in addressing some of these challenges. A cognitive radar possess the ability to learn from experience through continuous interactions with the environment. This ability is a key element of cognition, even in human beings. Thus, a cognitive radar is said to be "intelligent" in its functioning. This paper aims to explain the concept of cognitive radars, the different elements of cognition, and survey some of the latest literature on the design of cognitive radars for different applications.

2 Introduction

The cognitive radar was first conceived by Simon Haykin [1] in 2006. The idea is inspired from the echolocation system of a bat. A bat transmits sound waves into the environment and uses the reflected waves to determine the location of target objects. The bat is also capable of determining the velocity, elevation, size and other features of the target [2]. Moreover, it can adapt to the changing behavior of targets and successfully pursue them. It can also adapt to a changing environment. It is fascinating how these tasks are performed by the bat using a tiny brain and are yet a challenge to realise in physical radar systems. A major reason is that soon after birth, a bat learns the whole process through repeated interactions with the environment. This process of "learning from experience" is a key element of cognition and can be observed even in human beings. This notion of cognition can be used to make a radar system more "intelligent". Thus, unlike existing radar systems which work on a specific set of rules developed by a designer, a cognitive radar seeks to learn its own rules of behavior according to its environment. Since these learned ruled pertain specifically to the environment and scenario in which the radar works, they result in superior performance. Therefore, the presence of cognition in a radar can significantly improve its capabilities across a wide range of applications. It can also enable radars to perform multiple functions in different environments since the radar is capable of learning when introduced to a new environment.

The rest of the paper is structured as follows: First, the different types of radar systems are described and the key distinguishing factors of cognitive radars are noted. Then the term *cognition* is explored, followed by its mathematical description for practical applications. Finally, a basic cognitive radar system is explained after which a literature survey of the latest trends and techniques in cognitive radars is presented.

3 Types of Radars

Simon Haykin [3] distinguishes three classes of radars:

1. Traditional Active Radar (TAR): These radars act in a feed-forward manner (the receiver knows the transmitted waveform). Estimation algorithms might be used (e.g. Kalman Filter).

2. Fore-Active Radar (FAR): This type of radar contains feedback from the receiver to the transmitter. It is hence said to posses limited intelligence.

3. Cognitive Radar (CR): Develops rules of behavior in a self-organized manner through a process called learning from experience that results from continued interactions with the environment.

The TAR does not possess any learning capability and hence no cognition. The FAR however, is *adaptive* as a result of the feedback loop but its intelligence is limited because it cannot learn from experience on its own. The cognitive radar on the other hand, can learn from experience in a self-organized manner, much like a human beings or other animals. This ability is hence the major distinguishing factor of cognitive radars.

4 Elements of Cognition

In order to realise cognition in physical systems, it is imperative to understand what cognition is, and what it involves. In his conception of a cognitive radar, Haykin [3] uses Furster's [4] paradigm of cognition that consists of the following four fundamental building blocks

1. Perception-Action Cycle: The propagation of information from the receiver to the transmitter and the subsequent action performed by the transmitter based on this information. This forms a cycle where the transmitter acts, perceives the result of its actions and acts again accordingly (analogous to humans using their senses to perceive and act).

2. Memory: The recollection of past experiences. In physical terms, past experiences need to be stored and retrieved when necessary.

3. Attention: The algorithms that work on the perception-action cycle and memory.

4. Intelligence: The combination of the above three blocks to form the "intelligence" of the system.

5 A Basic Cognitive Radar

The fundamental blocks discussed in the previous section are put together to form a basic CR system, whose block diagram [1] is shown in Fig. 1. The transmitter transmits a signal into the environment and the radar returns are collected by the receiver. Other sensors provide additional information to the receiver about the state of the environment. These sensors help the receiver analyze the scene of the environment (e.g., humidity, temperature, etc). The receiver also contains

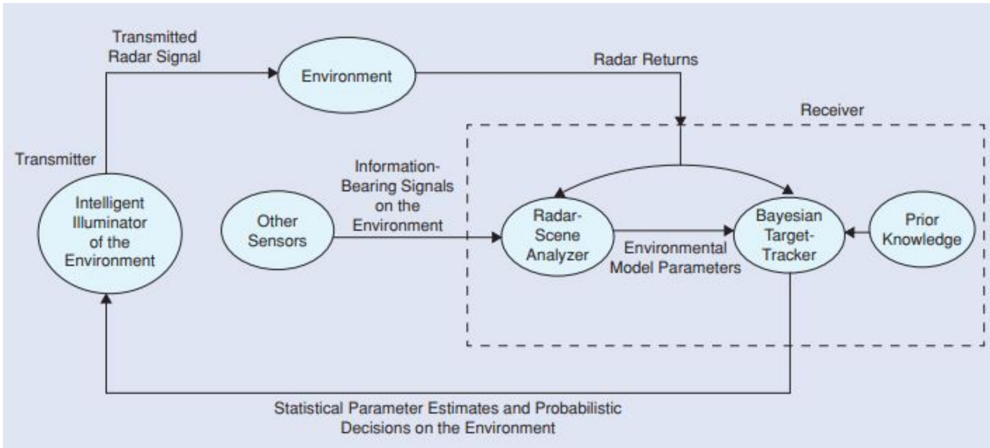


Figure 1: Block diagram of a CR system

prior-knowledge that might be useful to make decisions. For example, military radars could be provided prior knowledge in the form of a map of the surroundings. The radar might even update its knowledge during its interactions with the environment. The decision making system in the receiver takes into account both the analyzed scene and the prior knowledge along with the radar returns (the bayesian tracker is provided as an example). Post making the decision, relevant information is then fed back to the transmitter which adjusts its characteristics accordingly. In this sense, the transmitter controls the environment and receiver. It is hence said to be an intelligent illuminator of the environment.

6 Incorporating Cognition

Although a theoretical formulation of cognition exists, in order to embed it into physical systems a mathematical description is necessary. This involves breaking down the fundamental elements of cognition into mathematical formulations. The building blocks of cognition that were discussed in the previous sections can be realised in radar systems [3] as follows:

1. Perception-Action Cycle: The perception action cycle can be described using well developed engineering concepts such as:
   - Probabilistic modelling of perception at the receiver. This mainly includes bayesian methods.
   - Methods to make the transmitter optimally control the receiver through the environment. These include dynamic programming, reinforcement learning, etc.
   - Information theory to mathematically describe the link between transmitter and receiver.
2. Memory: The process of "learning from memory" is facilitated by memory in both the receiver and transmitter. The receiver might store information such as a map of the surroundings, sensor data from the scene analyzer, etc. The decision maker may also require information to be stored and updated such as in the case of priors when using bayesian methods.
3. Attention: The algorithms that exploit the perception-action cycle and memory to perform desired tasks form the attention block of the system. For example, the explore-exploit strategy often used in reinforcement learning.

4. Intelligence: The cohesive working of the three above blocks in the system is called the intelligence of the system.

7 A Literature Survey

Several approaches have been developed based on the ideas discussed in the previous sections to design cognitive radars for a variety of applications. Among these approaches, reinforcement learning (RL) proves to be the most promising. This is primarily because reinforcement learning enables the radar system to learn from experience through continuous interactions with the environment. A brief overview of reinforcement learning is provided in the subsequent section, followed by a review of few latest works on RL based design of cognitive radars.

7.1 Reinforcement Learning

Reinforcement learning (RL) is a technique wherein an agent learns behavioral rules through continuous interactions with the environment. This is analogous to how human beings, for example, learn to ride a bicycle. The entire process is modelled as a *Markov Decision Process* in which the agent transitions between *states* according to the *actions* that it performs. An interaction with the environment involves the agent performing a certain *action* which takes the agent to a specific *state*. The agent also receives a *reward* for its action. This *reward* serves as a kind of feedback for the agent regarding it's action i.e., favourable actions will result in higher rewards and unfavourable actions in lower rewards. By maximizing the rewards that it receives, the agent learns to perform the most favourable actions during its interaction with the environment. Hence, the reward system in the RL framework of a cognitive radar forms the feed-back loop. The transmitter is an "intelligent illuminator" in the sense that it uses these rewards (feed-back) to change its characteristics to meet a certain goal.

7.2 Reinforcement Learning based Beam-forming for Massive MIMO Radar Multi-target Detection

Mostafa *et al.* [5] propose a RL based method for beam-forming in Massive MIMO radar for multi-target detection. Beam-forming is a technique for focusing a signal in a particular direction. This is achieved by generating signals of particular phase, such that they constructively interfere and form a beam in the required direction. Beamforming not only helps to direct a beam in a particular direction, but it also optimizes the power usage of the transmitter by transmitting more power in the required direction and less power in the other directions.

The goal of the RL agent in this work is to learn the best way to beamform at the transmitter such that the probability of detection is improved, i.e., the transmitter is an "intelligent illuminator" of the environment. Thus, in this system, the transmitter is the major cognitive component. The RL framework for the system is described below:

- States: The total number of angle bins that contain targets. The number of states is thus determined by the total number of targets that the radar is designed to detect.

- Action: Beamforming according to the angle bins that contain targets. This takes the form of an optimization problem that seeks to optimize the beamforming matrix.

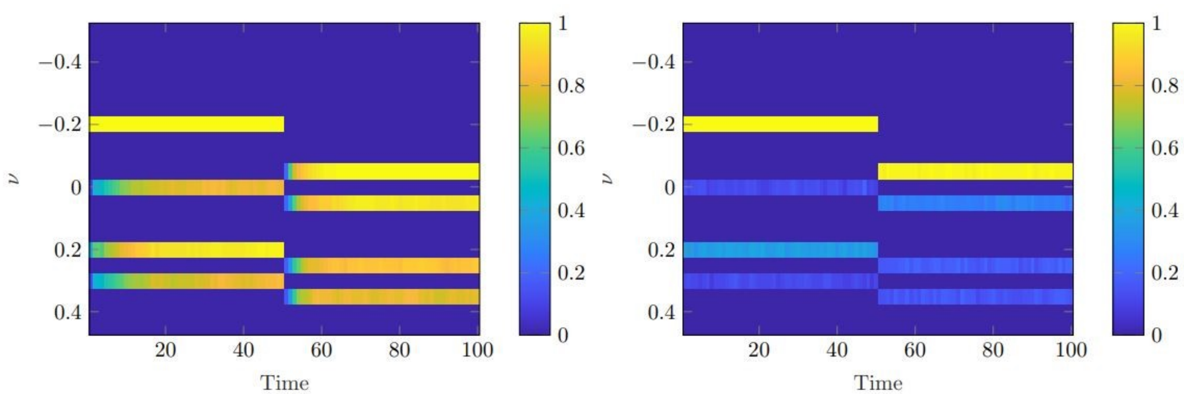
- Reward: The agent receives two kinds of rewards -

- Positive reward: The sum of the probability of detection (P_D) of the targets in the respective angle cells.
- Negative reward: The sum of P_D of the rest of the cells (that are not likely to contain targets).
- Total reward: Positive reward - Negative reward.

Recall that the purpose of the reward in the RL system is to give the agent an idea of which actions are favourable and subsequently the agent learns the best actions by maximizing its rewards. Thus, in the above mentioned reward framework, the agent maximizes the probability of detection by maximizing its rewards. The agent is trained using the Q-learning algorithm. In simple terms, the agent beamforms and performs detection using a certain signal model and detection theory which leads to a state transition, for which the agent receives the corresponding reward. Based on this reward, the agent decides the next action (facilitated by the learning algorithm) and undergoes the same process. Hence, after several trials, the agent learns the actions that give the best rewards.

Mostafa *et al.* show that their RL based beamforming method outperforms an omni-directional transmitter with equal power allocation at multi-target detection. A comparison of the two for dynamically changing environments is shown in Figure 2 and Figure 3. The first experiment was conducted by simulating four targets at spatial frequencies $-0.2, 0, 0.2, 0.3$, and the locations of the targets change at $T = 50$. The probability of false alarm is kept constant at 10^{-5}. In the second experiment, the SNR due to the targets decrease by 20% every 30 time units. The experiments use the signal model as defined in [6][7] and detection is performed by hypothesis testing using a robust Wald-type test. However, according to the application, the proposed RL framework can be used with a signal model and detection theory of the designer's choice.

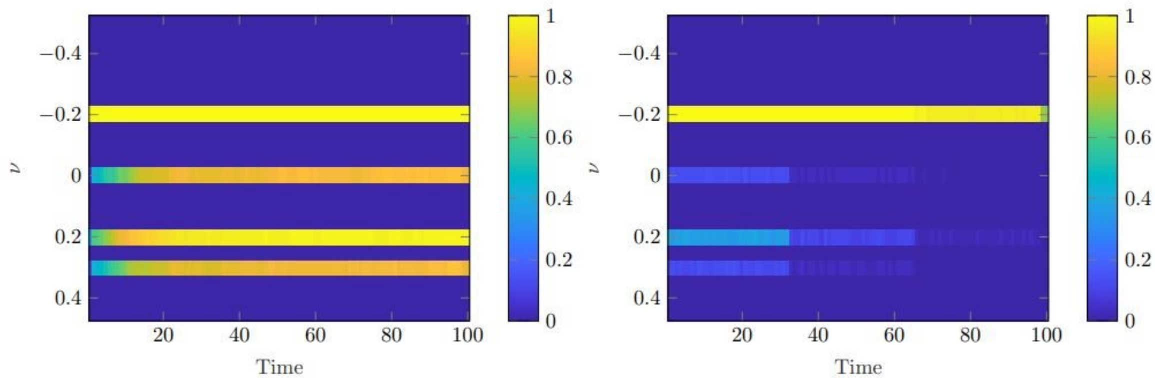
From Figure 2, the probability of detection (P_D) is observed to be significantly higher in the case of RL-based beamforming. However, the probability of detection for the RL agent is low at the beginning and as the agent learns, the detection performance increases. This can be seen in Figure 2a where the probability of detection increases to a high value in the first 10 time units. This is not the case when equal power is allocated since there is no learning involved. The Rl agent is also able to detect the targets at their new locations after $T = 50$. Again, the RL agent requires around 10 time units to learn.



(a) RL based beamforming

(b) Omnidirectional with equal power allocation

Figure 2: Performance comparison of omni-directional equal power allocation and RL based beamforming. The target location changes at $T = 50$.



(a) RL based beamforming

(b) Omnidirectional with equal power allocation

Figure 3: Performance comparison of omni-directional equal power allocation and RL based beamforming. The SNR due to the targets decrease by 20% every 30 time units.

In the results of the second experiment shown in Figure 3, they again find that the RL based system outperforms the equal power allocation system. The RL based system is less sensitive to decrease in SNR once it has learned the location of the targets. However, in he case of equal power allocation, the probability of detection significantly reduces due to the decrease in SNR of radar echos from a target. The authors also show that their RL based system outperforms the equal power allocation system at very low constant false alarm rates and very high antenna array resolutions.

Although only target detection is considered in this work, the superior performance of the RL based system in dynamic environments indicates the possibility of its success even for target tracking. Indeed, the application of RL to more complicated tasks such as tracking is currently of high interest in cognitive radar research.

7.3 Reinforcement Learning based Anti-jamming Frequency Hopping Strategies Design for Cognitive Radar

The performance of a radar can be severely affected when the receiver is overwhelmed with noise and this aspect of radars is sometimes exploited to disrupt its functioning. The deliberate disruption of a radar by saturating the receiver with noise or false information is called radar jamming. Jamming is often used during military operations to ensure that the enemy's radars cannot detect airplanes, tanks, etc.

In order to jam a radar, the jammer needs to know the radar's frequency band of operation. This is because it is not feasible in terms of power for a jammer to transmit noisy signals in all frequency bands (e.g., white noise). Moreover, the noise amplitude should be high enough to saturate the received radar echos and make detection impossible for the radar. But once the frequency band of operation of the radar is known, the jammer can transmit noisy signals in only these frequency bands to jam the radar. The scenario thus plays out as follows: the jammer constantly tries to figure out the frequency of operation of the radar and the radar in turn seeks to evade the jammer by trying to figure out (or avoid) the frequency of operation of the jammer. In other words, there is a continuous competition between the strategies of the jammer and the radar.

A naive strategy for the radar to avoid the jammer would be to randomly change its carrier frequency. This is called frequency hopping. However, if the radar could instead learn the jammer's strategy, it could optimize the use of its resources. This idea leads to the conception of a kind of cognition that involves intelligently avoiding jammers present in the radar's environment.

Kang *et al.* [8] tackle this problem using RL. They propose an RL based technique wherein the RL agent aims to learn/discover the jammer's strategy and intelligently hop the radar carrier frequency to avoid jamming. They experimentally show that the RL approach is significantly better than a random hopping strategy. Their RL framework is described as follows

- States: Each carrier frequency is a state. Hence, the number of states is determined by the set of carrier frequencies that the radar is capable of using.
- Action: A hop to a specific carrier frequency constitutes the action.
- Reward: The signal-to-interference-plus-noise ratio (SINR) is used as a reward. The frequencies at which the jammer is active will result in lower SINRs than the frequencies that are not jammed. The SINR can hence indicate the presence of a jammer.

Since the goal of the RL agent is to learn the actions that produce maximum rewards, the agent learns to hop the carrier frequencies to produce the best SINR, which consequently avoids the jammer. Simulations were performed to demonstrate the effectiveness of their method. A pulse wave radar with a coherent processing interval (CPI) of $N = 100$ pulses was considered where the carrier frequency is changed within a CPI. The radar's carrier frequency ranges from $3GHz$ to $5GHz$ in steps of $1MHz$. The simulation was performed for three methods —Q-learning based RL agent, DQN based RL agent and a random hopping strategy. Deep Q-network (DQN) [9] is another RL learning algorithm that combines the Q-learning method with a neural network.

The simulation results are shown in Figure 4. It is observed that the RL based methods have a significantly higher SINR than the random hopping strategy and hence result in better performance for the radar. This also indicates that the RL based radar has avoided the jammer better than the random hopping radar. Moreover, the RL agent that is trained using the DQN algorithm performs better than the agent trained using the Q-learning algorithm.

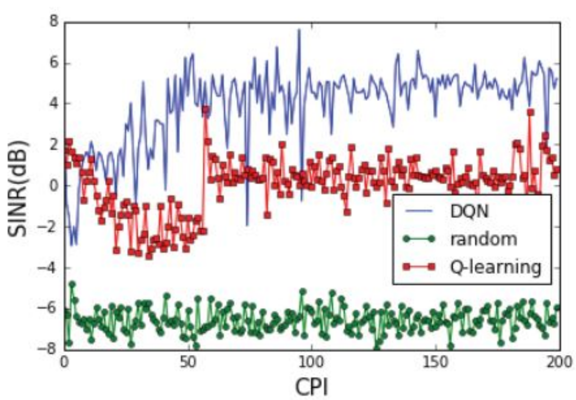


Figure 4: Performance comparison of the different methods

These results show the effectiveness of RL in avoiding jamming and ensuring security of radars. The cognitive element in this case is the ability of the radar to intelligently avoid jamming. Indeed, cognition could also be used to intelligently jam a radar! Wang *et al.* [10] propose such a method wherein the RL agent aims to maximize the jamming-plus-noise-to-signal ratio (JNSR). They show that their method is more effective in jamming radars than a random frequency hopping jamming strategy. Thus, cognition for security, particularly in military radars, plays an important role in both defensive and offensive situations.

7.4 A Machine Learning Radar Scheduling Method Based on the EST Algorithm

The functionality of radars is diverse, ranging from target detection and tracking in dynamic environments to imaging applications. Each of these functions involve multiple tasks to be processed. In certain applications like the military, radars are required to process several functions and their constituent tasks in real-time (within time constraints). Hence, such a radar must be able to optimize its resources to meet the application's demands. In this sense, the radar can be viewed as a real-time system. An important aspect of resource and time optimization is the scheduling of different tasks. Several algorithms such as the shortest job first, highest priority first, soonest completion time first, earliest start time first, etc., are commonly used in real-time systems to schedule different tasks. However, these algorithms operate according to a fixed set of rules that cannot be modified dynamically. This leads to the idea of making real-time systems schedule tasks dynamically like humans do, i.e., intelligently schedule tasks according to the situation.

Qu *et al.* [11] propose a RL method based on the earliest start time algorithm (EST) that improves the task scheduling performance when compared to the traditional EST algorithm, while maintaining a computational time that is suitable for real-time applications such as radars. They compute a total cost for scheduling a sequence of tasks and minimize this cost to find the optimal sequence. The total cost (J) of scheduling a sequence of N_{actual} tasks is defined below:

$$J = \sum_{n=1}^{N_{actual}} C(n) \tag{1}$$

Where $C(n)$ is the cost for an individual task n and is defined in a mean-squared error format:

$$C(n) = \frac{1}{N_{actual}} \left[p(n) \frac{t_{start}(n) - t_{res}(n)}{\tau} \right]^2 \tag{2}$$

Where $t_{start}(n)$ is the required start time of the task, $t_{res}(n)$ is the rescheduled start time as determined by the scheduler and $p(n)$ is the priority of the task n. The priority overcomes one of the drawbacks of the traditional EST algorithm which does not take into account the priority of tasks. τ is defined as follows:

$$\tau = |t_{earliest}(n) - t_{start}(n)|, \text{ where } t_{res}(n) < t_{start}(n) \tag{3}$$

$$\tau = |t_{latest}(n) - t_{start}(n)|, \text{ where } t_{res}(n) \geq t_{start}(n) \tag{4}$$

The ratio $\frac{t_{start}(n) - t_{res}(n)}{\tau}$ represents the deviation of the rescheduled time from the original start time. Thus, the cost incurred is higher as the rescheduled time is farther away from the required start time. Therefore, minimizing the cost will ensure least deviation of the rescheduled time from the original start time. The earliest time $t_{earliest}(n)$ is the earliest possible time that the task can be scheduled and the latest time $t_{latest}(n)$ is the latest possible time that the task can be scheduled, beyond which the task would be useless. Hence, if the scheduler cannot schedule the task within the time window $[t_{earliest}, t_{latest}]$ the task will be dropped. The process of scheduling the tasks is written as shown in Equation 5. The EST algorithm takes the set of tasks along with their required start times as input and returns a sequence of tasks with new scheduled start times.

$$t_{res}(n) = EST\left[t_{start}(n)\right] \tag{5}$$

The authors make the observation that shifting the start time of each task and then providing this as input to the EST algorithm will result in a differently scheduled sequence and a different total cost. Thus, adjusting the time shift t_{shift} of each task effectively can potentially reduce the total cost incurred.

RL is used to learn the adjustments that would result in the sequence of tasks that incur the least cost. In the RL framework, the set of t_{shift} are the states and switching to a new t_{shift} is an action. At each new state, the total cost J is evaluated. If J is lesser than the existing value, then the RL agent gets a positive reward of 1 unit. Otherwise, the agent is punished with a negative reward of 2 units. In this manner, the RL agent is able to learn the values of t_{shift} that can provide low total costs. However, this total cost may not be the global minimum. When the number of tasks to be scheduled increases, it becomes infeasible to find the global minimum for real-time applications and hence, a good and fast solution as proposed in this work is more suitable. Simulation results have shown that this method is 20 times better than the traditional EST scheduling algorithm.

8 Conclusion

The need to make machines intelligent is ever increasing in the modern world, and radars are no exception. This incorporation of intelligence in the form of cognitive capabilities has resulted in a new kind of radar called the cognitive radar. The cognitive radar as conceptualized by Simon Haykin formed a basic framework. Significant advances in the field of machine learning has provided a robust mathematical framework for cognition that can be used even for radars. It is hence not surprising that the design of cognitive radars is now one of the most active research topics in the radar domain. As indicated in this paper, reinforcement learning seems to be the most accurate mathematical framework that incorporates the notion of "learning from experience". It has been used to develop cognitive radars that apply intelligence in various aspects such as for beamforming, security, resource optimization, etc. Furthermore, there is immense scope in various other applications. Yet, cognitive radars of the present are intelligent only in specific aspects. The frontiers of cognitive radar would be the design of multi-functional radars, i.e., radars that can be used in different environments and perform various applications. Such radars could greatly help reduce the cost of deployment in several environments.

References

[1] S. Haykin. "Cognitive radar: a way of the future". In: *IEEE Signal Processing Magazine* 23.1 (2006), pp. 30–40. DOI: 10.1109/MSP.2006.1593335.

[2] C.F. Moss J.A. Thomas and M. Vater. *Echolocation in Bats and Dolphins.* Chicago, IL: Univ. of Chicago Press, 2004.

[3] Simon Haykin, Yanbo Xue, and Peyman Setoodeh. "Cognitive Radar: Step Toward Bridging the Gap Between Neuroscience and Engineering". In: *Proceedings of the IEEE* 100.11 (2012), pp. 3102–3130. DOI: 10.1109/JPROC.2012.2203089.

[4] Joaquín M. Fuster. *Cortex and Mind: Unifying Cognition.* Oxford, U.K.: Oxford Univ. Press, 2003. DOI: 10.1093/acprof:oso/9780195300840.001.0001.

[5] Aya Mostafa Ahmed et al. *A Reinforcement Learning based approach for Multi-target Detection in Massive MIMO radar.* 2021. arXiv: 2005.04708 [eess.SP].

[6] Jian Li and Petre Stoica. "MIMO Radar with Colocated Antennas". In: *IEEE Signal Processing Magazine* 24.5 (2007), pp. 106–114. DOI: 10.1109/MSP.2007.904812.

[7] Benjamin Friedlander. "On transmit beamforming for MIMO radar". In: *IEEE Transactions on Aerospace Electronic Systems* 48 (Oct. 2012), pp. 3376–3388. DOI: 10.1109/TAES.2012.6324717.

[8] Li Kang et al. "Reinforcement Learning based Anti-jamming Frequency Hopping Strategies Design for Cognitive Radar". In: *2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. 2018, pp. 1–5. DOI: 10.1109/ICSPCC.2018.8567751.

[9] V. Mnih et al. "Human-level control through deep reinforcement learning". In: *Nature* 518.7540 (Feb. 2015), pp. 529–533.

[10] Lulu Wang et al. "Optimal Jamming Frequency Selection for Cognitive Jammer based on Reinforcement Learning". In: *2019 IEEE 2nd International Conference on Information Communication and Signal Processing (ICICSP)*. 2019, pp. 39–43. DOI: 10.1109/ICICSP48821.2019.8958575.

[11] Zhen Qu, Zhen Ding, and Peter Moo. "A Machine Learning Radar Scheduling Method Based on the EST Algorithm". In: *2019 IEEE 18th International Conference on Cognitive Informatics Cognitive Computing (ICCI*CC)*. 2019, pp. 22–27. DOI: 10.1109/ICCICC46617.2019.9146101.