

Article

Collaborative Beamforming with DQN for Interference Mitigation in 5G and Beyond Networks

Alaelddin F. Y. Mohammed ¹, Salman Md Sultan ^{2,*} and Sakshi Patni ^{3,*}

¹ Information Technology, Department of International Studies, Dongshin University, 67, Dongshindae-gil, Naju-si 58245, Republic of Korea; alaelddin@dsu.ac.kr

² Innovative Skills LTD, 693, West Shewrapara, Mirpur, Dhaka 1216, Bangladesh

³ Department of Computing, Gachon University, Seongnam-si 13120, Republic of Korea

* Correspondence: salmanmdsultan92@gmail.com (S.M.S.); drsakshi5@gachon.ac.kr (S.P.)

Abstract: This paper addresses the problem of side lobe interference in 5G networks by proposing a unique collaborative beamforming strategy based on Deep Q-Network (DQN) reinforcement learning. Our method, which operates in the sub-6 GHz band, maximizes beam steering and power management by using a two-antenna system with DQN-controlled phase shifters. We provide an OFDM cellular network environment where inter-cell interference is managed while many base stations serve randomly dispersed customers. In order to reduce interference strength and improve signal-to-interference-plus-noise ratio (SINR), the DQN agent learns to modify the interference angle. Our model integrates experience replay memory with a long short-term memory (LSTM) recurrent neural network for time series prediction to enhance learning stability. The outcomes of our simulations show that our suggested DQN approach works noticeably better than current DQN and Q-learning methods. In particular, our technique reaches a maximum of 29.18 dB and a minimum of 5.15 dB, whereas the other approaches only manage 0.77–27.04 dB. Additionally, we significantly decreased the average interference level to 5.42 dB compared to competing approaches of 38.84 dB and 34.12 dB. The average sum-rate capacity is also increased to 3.90 by the suggested strategy, outperforming previous approaches. These findings demonstrate how well our cooperative beamforming method reduces interference and improves overall network performance in 5G systems.

Keywords: collaborative beamforming; Deep Q-Network (DQN); 5G networks; interference mitigation; LSTM neural networks; OFDM; sub-6 GHz



Citation: Alaelddin, F.Y.M.; Sultan, M.S.; Patni, S. Collaborative Beamforming with DQN for Interference Mitigation in 5G and Beyond Networks. *Telecom* **2024**, *5*, 1192–1204. <https://doi.org/10.3390/telecom5040060>

Academic Editor: Sotirios K. Goudos

Received: 20 September 2024

Revised: 28 October 2024

Accepted: 20 November 2024

Published: 3 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The exponential growth in mobile data usage opened the door for the deployment of 5G and beyond networks. In 5G, it is important to ensure reliable services for the user because users demand multi-tasking in their smartphones, including video streaming, voice calls, the usage of data, and so forth. However, in order to serve the large number of users with high data rates, we need to operate our signal at higher frequencies. Specifically, we currently use sub-6 GHz frequency to serve the highest data rate in 5G networks [1]. However, the higher frequency can create a propagation loss so that the gain of an antenna-based base station (BS) can be low. Beamforming is one of the signal processing-based spatial filter techniques that can reduce loss by creating higher antenna gain [2]. The beamforming technique is widely used in 5G networks to create directional energy towards a particular user [3]. Multiple antennas (2, 4, 6, 8, 16, 32, 64) can be used to generate the beamforming. Nevertheless, beamforming has some issues, especially two antenna beamforming, such as higher side lobe due to large antenna spacing [4]. Figure 1 illustrates the different radiation patterns where the x-axis and y-axis represent the degree and power of the beam for three different antenna element distances.

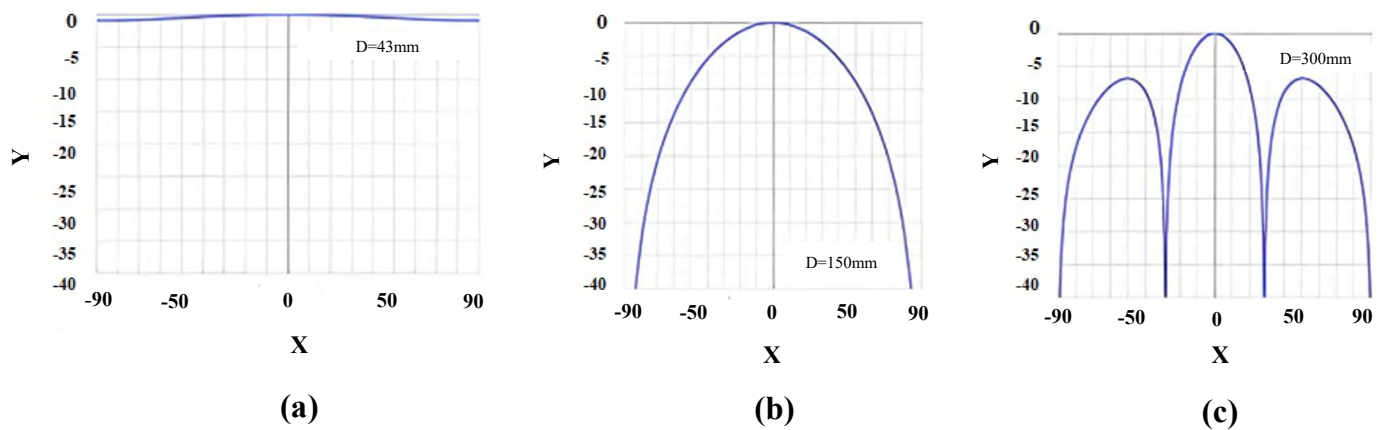


Figure 1. (a) Radiation pattern with antenna spacing $d = 43$ mm. (b) Radiation pattern with antenna spacing $d = 150$ mm. (c) Radiation pattern with antenna spacing $d = 300$ mm.

Here, we can observe that the beam width is wider because of the small antenna distance ($d = 43$ mm). Moreover, the directivity of the main lobe is narrower after increasing the antenna spacing ($d = 300$ mm), but the higher side lobe arrives in Figure 1c, which can create interference to another user where multiple BS exist near to one another. In 5G beamforming, we need to ensure good data transmission quality for the user, so we need to control the side lobe to reduce interference. Collaborative beamforming is one of the promising techniques to control the side lobe [5]. Collaborative beamforming creates a virtual antenna array where multiple independent base stations are located to cooperate with one another [6]. The collaborative beamforming enhances the voice quality during transmission, such as the signal interference noise ratio (SINR), with a minimum amount of power. There are many collaborative approaches, such as power control and beam-steer control, available to deal with the higher side lobe problem [1,6]. The authors in [6] studied the excitation current control using hybrid discrete and continuous optimization to reduce the side lobe for enhancing the energy quality. The authors in [1] proposed a Deep Q-Network (DQN) based joint power control technique to reduce the interference power for better sum-rate capacity in the uniform linear array (ULA) based beamforming sector. However, the authors improved the voice and data bearer's signal quality by increasing the number of antennas up to 32, which basically increases the cost of the production. Therefore, we need an optimized solution to overcome cost by minimizing the number of antennas. In this work, we propose a two-antenna collaborative beamforming technique using deep reinforcement learning to improve the SINR and sum-rate capacity by controlling beam-steering and the power of the interfering signal [7]. The following are the detailed contributions to this paper:

1. We introduce the cost-effective DQN solution to control the beam-steering to reduce the slide lobe interference.
2. We mainly focus on reducing the interference power controlling beam-steering using the proposed DQN.
3. In this work, our proposed solution applies to a two-antenna collaborative environment to reduce the cost of operation.
4. The proposed DQN agent can decide a specific shift angle (i.e., right or left shift angle) depending on the time delay in beam-steering to reduce the interference's power and increase the serving power and SINR.

The remainder of this paper is organized as follows. The details of the related work are provided in Section 2. Section 3 presents our proposed DQN-based 5G beamforming architecture. We show the performance evaluation and comparison in Section 4 and draw the conclusion in Section 5.

2. Related Work

Beamforming is a signal processing technique that can be optimized using machine learning approaches [8]. By leveraging neural networks, reinforcement learning, and other algorithms, machine learning enables beamforming systems to intelligently adapt beams based on dynamic environments. Key applications of machine learning include beam pattern prediction, adaptive beam steering, direction of arrival estimation, and beam selection. With machine learning, beamforming systems can achieve more flexible, intelligent, and data-driven optimization of antenna arrays beyond traditional approaches. Several recent studies have explored hybrid beamforming/combining for millimeter wave (mmWave) multiple-input multiple-output systems to achieve a balance between performance and cost [9]. Table 1 summarizes the related work.

Table 1. Related work summary.

Topic	Work	Contribution	Pros	Cons
Machine Learning for Beamforming Optimization	[8–11]	Machine learning techniques enhance beamforming in mmWave MIMO systems by optimizing beam pattern prediction, selection, and user localization.	Flexible and adaptive optimization in dynamic environments [8]; reduced complexity in real-time processing [9].	High computational cost and complexity in training large datasets [10]; limited generalizability across varied configurations [11].
Hybrid Beamforming Systems	[10,12]	Development of neural hybrid beamforming approaches using autoencoders for flexible architecture compatibility and fast convergence.	Improved flexibility in beamforming/combining architectures [10]; fast training convergence [12].	Dependence on specific architecture configurations limits general application [10].
Reinforcement Learning-Based Beamforming	[9,13–15]	Deep reinforcement learning (DRL) applications in beamforming for dynamic adaptation in MIMO and V2X communications.	DRL enables high adaptability in changing environments [9]; enhanced performance in vehicular communications [15].	Training is resource-intensive, requiring high computational power [13].
Beam Scheduling and Localization	[16,17]	Adaptive beam scheduling and localization techniques in distributed MIMO systems using machine learning for optimized performance.	Effective in large-scale MIMO and subleased RF slices [16]; improved accuracy in localization under diverse channel conditions [17].	Susceptible to performance degradation in highly dynamic environments [16].
Physical Layer Security	[18,19]	Combination of beamforming and IRS for enhanced security in IoT networks and cooperative jamming networks.	Enhanced security against eavesdropping and jamming [18]; efficient energy utilization in IRS-assisted networks [19].	Complexity in coordinating IRS phase shifts and beamforming in real time [19].
Collaborative and Distributed Beamforming	[20–22]	Collaborative strategies for MIMO beamforming and dynamic, distributed algorithms for coordinated beamforming.	Distributed approach reduces the need for centralized CSI [21]; effective in multiuser and large MIMO scenarios [20].	Limited by inter-user interference and scalability challenges [22].
Energy Efficiency and UAV Beamforming	[6,14]	Optimization of energy consumption in UAV collaborative beamforming and sidelobe reduction in sensor networks.	Reduced energy consumption in UAV beamforming [14]; effective sidelobe management [6].	Limited scalability in multi-UAV scenarios; high power requirements for large sensor networks [6].

Authors in [10] addressed this gap by proposing a neural hybrid beamforming/combining (NHB) system utilizing autoencoders. NHB leverages a machine learning-based design applicable to various beamforming/combining architectures, offering flexibility and compatibility not found in prior methods. Furthermore, the proposed iterative training strategy ensures fast convergence, making NHB a promising solution for practical mmWave MIMO systems. The work in [13] focused on leveraging machine learning (ML) techniques to optimize beamforming configurations and overcome challenges like complex channels, hardware constraints, and high complexity. The paper reviews existing ML methods, discusses their applications and potential benefits, explores future research directions, and presents a case study comparing ML-based and conventional approaches. In essence, it highlights how ML can be used to unlock efficient and intelligent beamforming for improved mmWave communication. Research conducted in [4] developed a recurrent neural network model for predicting optimal beam patterns in mmWave systems. The research in [5] applied Q-learning with received signal strength (RSS) feedback to adaptively adjust beams in indoor 60 GHz networks. These studies demonstrate the potential of machine learning to enable intelligent beamforming by modeling high-dimensional spaces and optimizing based on data. In addition, the study in [12] developed a neural hybrid beamforming approach compatible with different system configurations. The method trains an autoencoder model in an unsupervised manner to obtain the analog and digital beamforming parameters. This demonstrates the potential of leveraging artificial intelligence for flexible hybrid beamforming optimization in mmWave MIMO systems.

Furthermore, the approach in [16] proposed a novel machine learning-based beam scheduling approach specifically designed for wireless virtualization with large-scale MIMO and subleased radio frequency slices. This approach leverages ML for optimized scheduling, aiming to outperform existing methods. Another approach, in [17], proposed an adaptive BeamMaP model for user localization in distributed massive MIMO systems. The approach leverages training data across various channel conditions and adaptively steers beams based on estimated user locations. This demonstrates the potential of adaptive beamforming techniques to enable robust localization in challenging outdoor settings with multipath and dynamic weather conditions. Aspects of secure communications in Internet of Things networks have been examined in earlier research, such as the use of intelligent reflecting surfaces and cooperative jamming. Nevertheless, there is still a lot to learn about how to combine edge computing with deep reinforcement learning, namely deep deterministic policy gradient (DDPG), to maximize energy economy and confidentiality in IRS-assisted cooperative networks [18].

To improve wireless communication security, The work discussed in [19] investigated the combination of physical layer security methods with intelligent reflecting surfaces. Nevertheless, there is more work to be done in the area of jointly optimizing transmit beamforming and intelligent reflecting surfaces (IRS) phase shifts for energy efficiency maximization under poor channel state information (CSI), especially when using deep reinforcement learning in an anti-jamming setting. On the other hand, the investigation in [20] proposed that the beam codebook design for multiuser MIMO systems has traditionally depended on either reference signal received power (RSRP)-based user clustering or complete CSI feedback, which often leads to resource-intensive operations or restricted accuracy. Though mostly limited to offline training and fixed beam topologies, recent machine learning-based techniques have shown promise and opened the door for more flexible and effective solutions in dynamic contexts.

Conventional coordinated beamforming techniques in large MIMO systems have mostly depended on centralized optimization methods, which have a high computing cost and need global CSI, making them unsuitable for use in dynamic mobile situations by [21]. Although machine learning approaches for beamforming have been studied recently, there is still a significant problem in creating low-complexity, distributed algorithms that can adapt to mobile circumstances and yet achieve near-optimal performance. To solve the multi-objective optimization issue of unmanned aerial vehicle (UAV) enabled

collaborative beamforming in air-to-ground communication systems, The findings in [14] devised a multi-agent deep reinforcement learning method. By optimizing UAV locations and excitation current weights, their suggested HATRPO-UCB algorithm simultaneously maximizes transmission rate and minimizes energy consumption, outperforming conventional techniques in dynamic situations. Using tensor decomposition and a channel, frequency, and temporal attention (CFTA) method, the authors in [15] created an attention-deep reinforcement learning strategy for joint beamforming optimization in reconfigurable intelligent surface (RIS) assisted mmWave massive MIMO vehicle-to-everything (V2X) communications. By combining tensor-based channel modeling with DRL-based optimization of RIS phase shift matrices and BS beamforming, their suggested approach tackles the problems of fast time-varying propagation and dynamic vehicular networks, showing enhanced performance for cooperative automated driving applications. The study in [23] developed an improved adaptive model for MIMO systems. They collected multipath RSS training data across line-of-sight and non-line-of-sight conditions. Using these data, their model adaptively steers beams based on estimated user locations to improve localization accuracy. Simulations demonstrate reduced root mean square error close to Bayesian bounds, outperforming support vector machine (SVM) and K nearest neighbor (KNN) approaches, proving the effectiveness of adaptive beamforming for robust 5G outdoor localization.

Further research should actively explore the collaborative development of learning techniques tailored for beamforming across different antenna array configurations. Although many promising approaches exist [11,24,25], beamforming techniques would benefit from continued collaborative efforts to pursue further improvements. However, finding the optimal method for practical scenarios with low-resolution phase shifters and ADCs remains an open challenge. Existing approaches often rely on specific system configurations and lack generalizability across different architectures [22]. Thus, there is still room for further development in this area.

3. Proposed DQN-Based 5G Beamforming Architecture

In Figure 2, the proposed 5G beamforming architecture is presented where multiple base stations BS1, BS2 BS_n are randomly allocated independently. In this work, we consider the orthogonal frequency-division multiplexing (OFDM) cellular network as similar to [1]. In the network, the green and red signal indicate the served beamforming and interference to a particular user. The users are randomly distributed in the associated area. Each BS is supposed to have two antennas, which are used to serve as voice bearers at sub-6Ghz frequency bands with the association of one DQN phase shifter. The objective of the DQN phase shifter is to control the angle of the interference. In Figure 2a we can observe that BS1, BS2, and BS3 serve independently to users U1, U2, and U3, respectively, and in the meantime, the interference attacks with angle θ_{it} from BS2 and BS3 to U1 and U2, respectively. In Figure 2b, our proposed DQN controls the steer with angle θ_{it+1} of the interference from BS2 and BS3 in such a way that the power of the interference can be reduced as well as the directivity of the serving beamforming being almost unchanged. Figure 3 shows the effects of changing the angle in terms of radiation pattern, where the x-axis and y-axis represent the degree and power of the beam. As we can observe from Figure 3, the interference first arrives at angle θ_{it} .

Then, after changing the phase of θ_{it} , it goes to θ_{it+1} and reduces the directivity approximately -2 dB. However, the directivity of the main beam remains the same even when we steer the angle of the interference signal. Basically, the beam-steering array factor (BSF) produces the steering angle of the main beam to serve a particular user. Similar to [1], we determine the beam steering array factor, $BSF(\theta_k)$, as in (1).

$$BSF(\theta_k) = \frac{1}{\sqrt{N}} \sum_{k=1}^N e^{jN \frac{\pi}{\omega} d \cos(\theta_k)} \quad (1)$$

where N = number of antenna, ω = wavelength, θ = angle of arriving signal, and d = antenna space. As we have mentioned earlier, our proposed DQN controls the angle θ_{it} of the particular interference signal so that interference power can be reduced. Therefore, the steer of interference signal is followed as in (1). After calculating the $\text{BSF}(\theta_k)$, we can derive the received power and interference power in terms of downlink transmission from BS to a particular user by (2) and (3) as per [1]:

$$R_p = P_s \times h_s \times \text{BSF}(\theta_k), \quad (2)$$

$$I_p = P_i \times h_i \times \text{BSF}(\theta_k), \quad (3)$$

where R_p is the received power, I_p is the interference power, P_s is the serving power of a particular base station (i.e.; BS1), I_p is the interference power of a particular base station (i.e., BS2), and h_s and h_i are the channel gain of a particular BS (i.e., BS1, BS2 respectively). Now, we can calculate the SINR [1] for the served user at a particular time t in (4).

$$\gamma(t) = 10 \log_{10} \left(\frac{R_p}{I_p + N_p} \right), \quad (4)$$

where N_p = noise power. In this paper, we apply our DQN to control the phase shifter to get the better SINR, $\gamma(t)$ in terms of downlink voice bearer transmission.

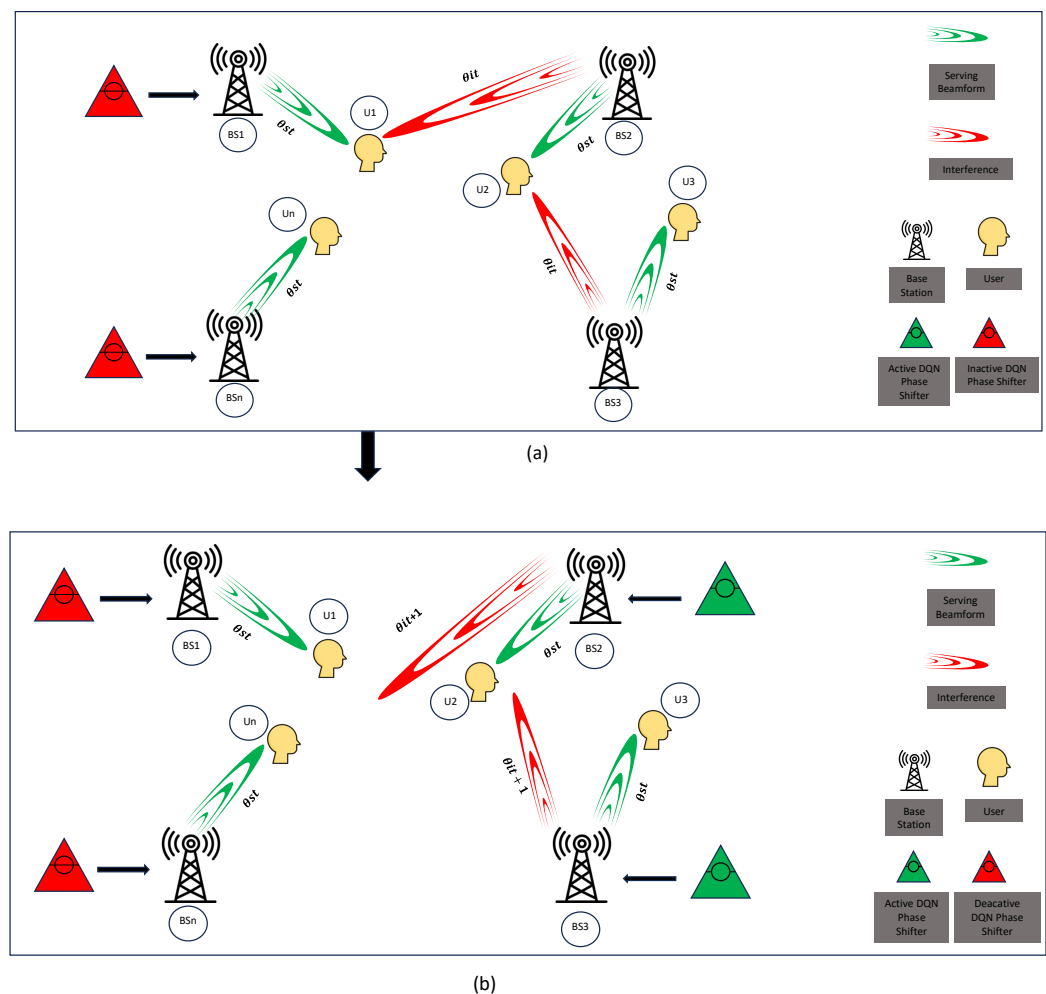


Figure 2. The proposed DQN-based 5G beamforming architecture: (a) Beamforming and interference in a traditional 5G network without DQN control, (b) Proposed DQN-based 5G beamforming with interference angle steering.

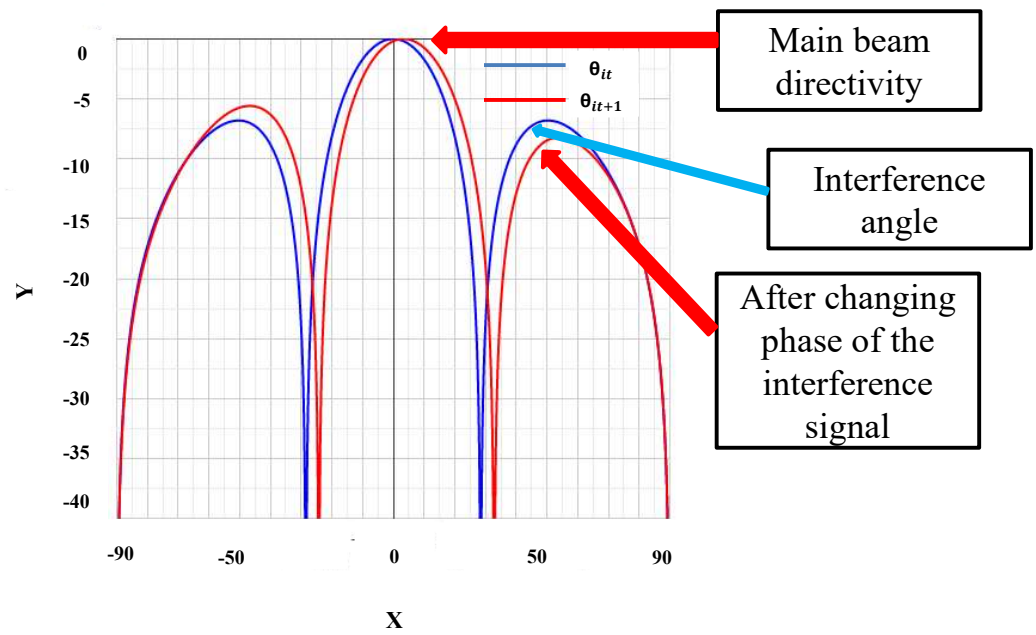


Figure 3. Radiation pattern of different phase angles.

Reinforcement Learning-Based 5G Beamforming Algorithm

In this paper, we introduce our DQN-based deep reinforcement learning, which can control the phase shifter to reduce the side lobe interference. In our case, the phase shifter is called a DQN agent, which is responsible for taking the best action α_t from a state s_t in our cellular OFDM-based environment. The DQN phase shifter's objective is to steer the interference angle, which is associated with reducing the interference power and enhancing the SINR. In our work, the state, s_t , is organized with different information such as coordinates of users, coordinates of the base station, angle of the signal, serving power, and interference power from a particular base station for the user. The proposed DQN phase shifter can choose only one action α_t among two possible actions, such as left shift angle and right shift angle. According to the concept of time delay in beam steering (if the time delay increases the phase moves to the left and, otherwise, the phase shifts to the right), we assume that if the served user is in the right direction from the interference signal the phase shift goes in the right direction; otherwise, it goes in the left direction. Therefore, the DQN phase shifter can either decrease the time delay or increase the time delay. We select the value of shift at 0.001 s, similar to [1] for either decreasing or increasing time delay. For each period of phase shifting, the interference signal loses some power, and the server signal's strength can be increased. Initially, the agent explores the environment by taking random actions because the agent does not know anything about the environment. After exploring, the agent exploits the environment by taking action from the neural network model. This method is called the epsilon-greedy technique.

After taking action α_t , the DQN agent takes reward as the difference between served power P_s and interference's power I_p . If our proposed DQN predicts the action properly, then the served power will be stronger than the interference power; otherwise, the interference power will be higher. Having this feedback from the environment to our proposed DQN agent can improve the predicted action. In this term, the reward is calculated in (5).

$$r_t = P_s - I_p \quad (5)$$

In the proposed method, our objective is to fulfill the SINR requirement for voice downlink transmission. The requirement is that our system must reach $\text{SINR} \geq 3$ dB. If the DQN does not fulfill this requirement, the system fails to achieve the goal. During the training session, an essential technique is introduced, called experience replay memory, to reduce the correlated data between the action value and the target value,

$Y = r_t + df \cdot \text{argmax}(Q(\text{nextstate}, \text{nextaction}))$, where df is discount factor that is between 0 and 1. This technique acts as the trajectory of the current state s_t , action a_t , reward r_t , and next observation s_{t+1} and stores the experience (s_t, a_t, r_t, s_{t+1}) into a replay memory (M) in each time step to reuse the past experience for better and stable performance after sampling is completed uniformly from the replay memory. Moreover, DQN uses the mini-batch technique to sample data randomly from the memory to train the neural network model via stochastic gradient descent so that the data are independent of each other. Thus, experience replay memory reduces the correlation between datasets. The main goal of the DQN is to optimize the loss function $L(\theta)$ of the neural network, as in (6).

$$L(\theta) = (Y - \theta)^2. \quad (6)$$

In our DQN, we use a long short-term memory (LSTM) recurrent neural network model that estimates the output based on time series data. Figure 4 shows the flowchart of our proposed DQN system, and Figure 5 presents the proposed LSTM model summary with a fully connected neural network, where our first layer contains LSTM with eight units and the remaining layers built with a Dense layer with four and two units, respectively. Note that the final Dense layer's unit is two because our proposed DQN agent predicts binary action (i.e., left shift or right shift).

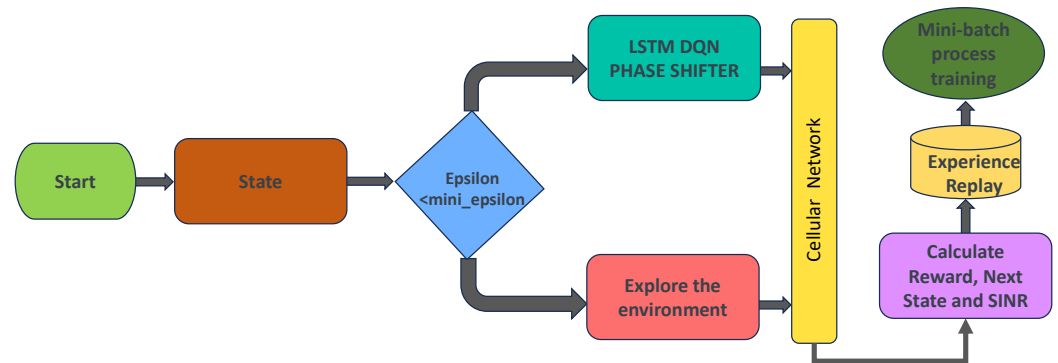


Figure 4. Flow chart of proposed DQN model.

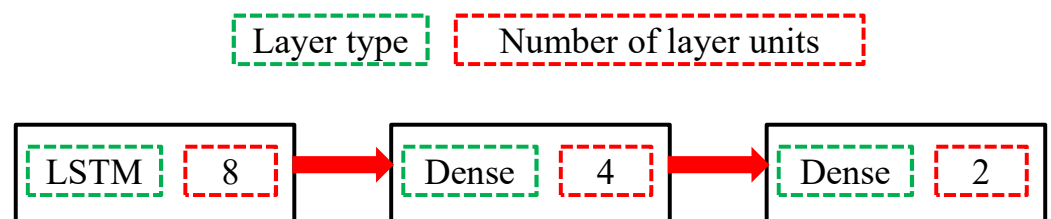


Figure 5. Model summary of proposed LSTM + Dense.

4. Performance Evaluation

In our simulation, we develop the environment using Python language. Python is an open-source language that helps to build large environments in a very fast way [26]. We also use the TensorFlow framework [27] to train our DQN model under the Anaconda environment [28]. The layer structure is easily made from Keras library [29]. To make our environment as well as the DQN model, we use some parameters that are given in Table 2. The parameter of Table 3 helps to fine tune our DQN model appropriately by a trial and error method.

Signal to interference plus noise ratio (SINR) is an essential metric that represents how well the signal quality is received by a receiver. It indicates how strong the desired signal is compared to the combined effect of interference and noise. A higher SINR typically indicates better signal quality and consequently better communication performance. In our proposed DQN-based method, the SINR at each successful step is higher than in other

existing work [1]. Higher SINR indicates better quality of service and communication. The proposed DQN (blue line) in Figure 6 is constantly higher than the other two existing methods [1]. This is also an indication that our SINR achieved through the proposed method provides better quality of service (QoS).

Table 2. Environmental parameters.

Parameters	Values
Cellular environment type	Circular
Maximum transmit power of BS	40
Bearer	Voice
Minimum SINR	3 dB
Number of antennas, N	2
Transmission type	downlink
Frequency band	2.1 GHz
Cell radius, r	350 m
Inter-site Ddistance	525 m

Table 3. Training parameters.

Parameters	Values
Discount factor, Size of state	8
Size of action	2
Minimum exploration rate	0.001
Learning rate	0.001
Maximum size of memory	100
Loss function	Categorical_Crossentropy
Activation function in hidden layer	Relu
Activation function in output layer	Sigmoid
LSTM model type	Sequential

In Figure 6, we can see that our proposed technique minimum SINR is 5 dB, whereas the existing DQN and Q-Table give the SINR near one and two, respectively. Moreover, our proposed DQN holds a very good SINR, which is greater than 20 dB, whereas others cannot cross 20 dB SINR. Therefore, the proposed DQN improves significantly compared to the existing solution. Furthermore, we derived a comparison in Table 4 according to serving power, interference power, SINR, and sum-rate capacity. The sum-rate capacity is calculated (similar to [1]) as in (7).

$$\text{Sum-Rate Capacity} = \sum_{i=1}^T \log_2(1 + \gamma(t)). \quad (7)$$

In Table 4, we can observe that our proposed DQN uses minimum serve power of 8.51 dB whereas the maximum served does not cross 20 dB. However, the served power of the existing technique [1] is greater than our proposed technique. Moreover, the existing DQN and Q-Table techniques used 38.84 and 34.12 as average interference power, which are very high compared to our proposed DQN, which is only 5.42. Due to having very low average interference power by controlling beam steering intelligently via our proposed method, the served power is higher than interference power; as a result, our SINR achievement is much better than the existing DQN and existing Q-Table method. Finally, our proposed DQN technique shows an average sum-rate capacity of 3.90, whereas existing DQN and existing Q-Table are 3.07 and 3.45, respectively.

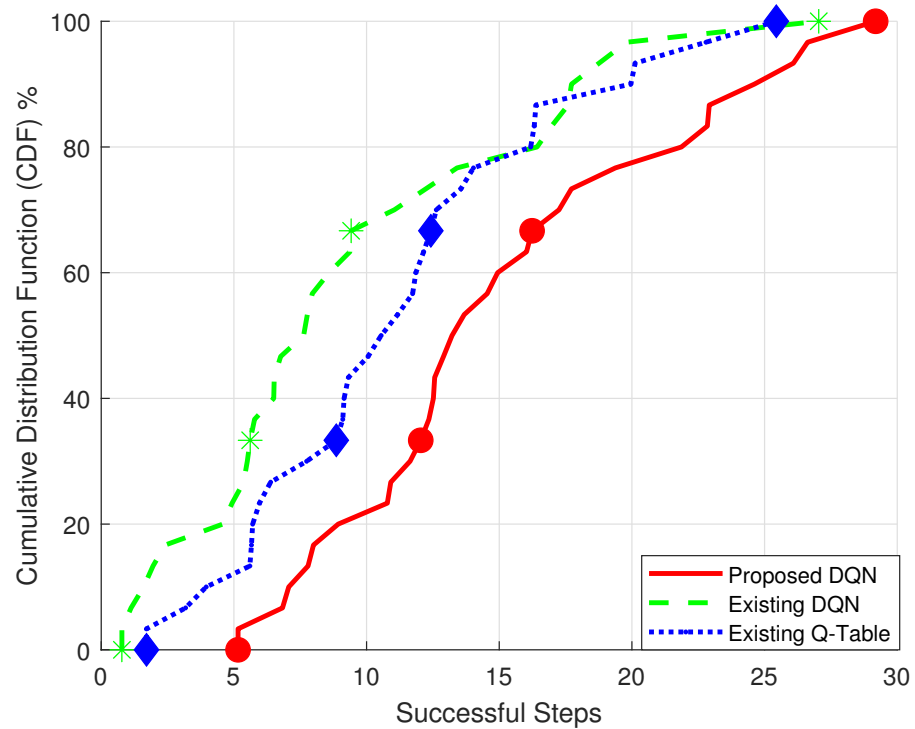


Figure 6. SINR at each successful step.

Table 4. Comparison table between proposed DQN, existing DQN, and existing Q-Table.

Parameter	Existing DQN [1]	Existing Q-Table [1]	Proposed DQN
Minimum Serve Power	28.005	23.707	8.51
Maximum Serve Power	46.32	47.34	17.60
Average Serve Power	41.4	35.5	15.73
Minimum Interference Power	27.09	22.61	0.78
Maximum Interference Power	42.2	45.8	7.9
Average Interference Power	38.84	34.12	5.42
Minimum SINR	0.77	1.70	5.15
Maximum SINR	27.04	25.44	29.18
Average SINR	9.37	11.32	16.44
Average Sum-Rate Capacity	3.07	3.45	3.90

Interference power refers to the unwanted signals or noise that degrade the quality of the received signal. Interference can arise from various sources and can have a significant impact on the performance of the communication signal. Managing the interference power is crucial for maintaining the reliability, capacity, and quality of service in 5G networks. The lower the interference power, the more reliable the signal would be. With our proposed DQN architecture method, it can be seen in Figure 7 that interference power is very low compared to the other existing methods [1] by controlling phase shifters. Thus, lower interference power indicates better communication between BS and U.

In our proposed DQN-based architecture, we also reduced the served power in each step compared to [1], as we can see in Figure 8. To control the beam forming and provide the best service with less power is highly cost-effective. In our proposed method, we make sure by taking sub-optimal action that at each step that the serving power from the base station is low while providing good service.

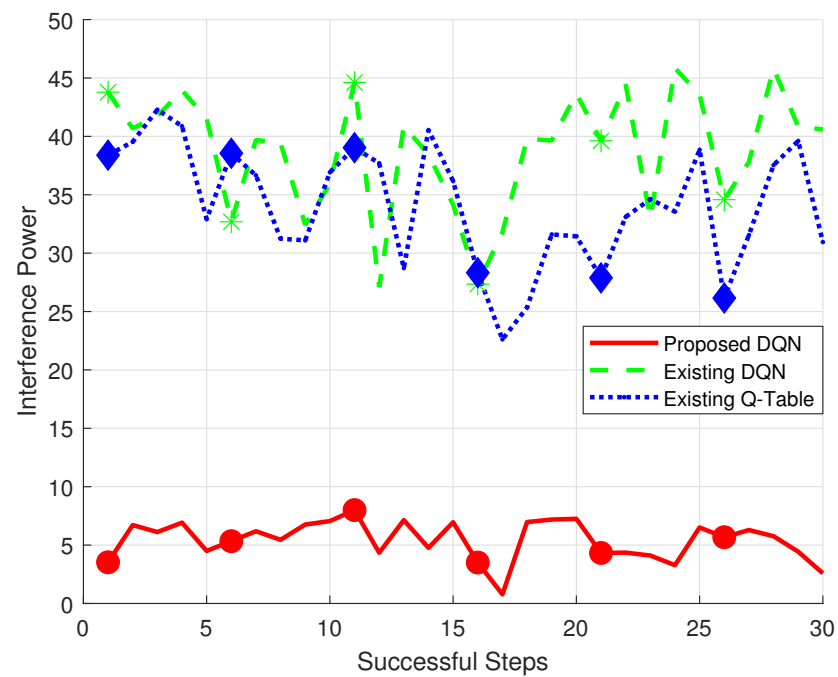


Figure 7. Interference power vs. successful step.

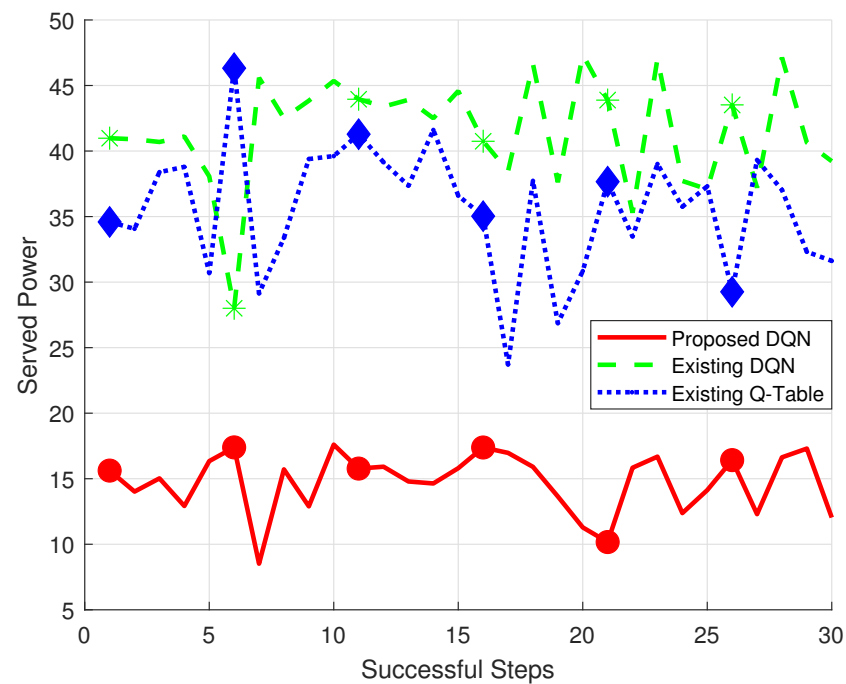


Figure 8. Served power comparison at each step.

5. Conclusions

This article addressed the essential issue of side lobe interference in 5G sub-6 GHz networks using a unique collaborative beamforming strategy based on DQN reinforcement learning. Maintaining high-quality service becomes more dependent on efficient interference control as network densification rises to meet increasing bandwidth demands. In order to improve learning stability, our suggested method integrates experience replay memory and LSTM networks for time-series prediction into a two-antenna system with DQN-controlled phase shifters. With many base stations servicing randomly dispersed customers in an OFDM cellular network environment, this strategy proved to be much

better than previous approaches. Improved average sum-rate capacity (3.90), decreased average interference power (5.42 dB), higher SINR performance (5.15–29.18 dB range), and lower served power levels (average 15.73 dB) are among the major accomplishments. These findings demonstrate that, by reducing interference while preserving signal quality, our strategy may greatly improve 5G network performance, beating both conventional DQN and Q-Table techniques. Our method's ability to adapt to changing network circumstances shows promise for increasing spectrum efficiency in dense deployments. Future research might examine issues with practical implementation, scalability in bigger networks, and application to mmWave frequencies. With its more effective and flexible beamforming techniques for 5G wireless communications, this study offers a viable avenue for 5G and beyond network optimization.

Author Contributions: A.F.Y.M.: Conceptualization, methodology, formal analysis, validation, writing original draft; S.M.S.: Conceptualization, writing original draft; S.P.: Formal analysis, validation, visualization, review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

Conflicts of Interest: Author Mr Salman Md Sultan was employed by Innovative Skills LTD. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Mismar, F.B.; Evans, B.L.; Alkhateeb, A. Deep reinforcement learning for 5G networks: Joint beamforming, power control, and interference coordination. *IEEE Trans. Commun.* **2019**, *68*, 1581–1592. [\[CrossRef\]](#)
2. Van Veen, B.D.; Buckley, K.M. Beamforming: A versatile approach to spatial filtering. *IEEE Assp Mag.* **1988**, *5*, 4–24. [\[CrossRef\]](#)
3. Cheng, B.; Huang, X.; Zhang, K.; Wu, F.; Leng, S. Cooperative Caching and Beamforming for CoMP-Enabled Mobile Edge Networks. In Proceedings of the 2022 IEEE 22nd International Conference on Communication Technology (ICCT), Nanjing, China, 11–14 November 2022; pp. 1038–1043. [\[CrossRef\]](#)
4. Khunteta, S.; Chavva, A.K.R. Recurrent Neural Network Based Beam Prediction for Millimeter-Wave 5G Systems. In Proceedings of the 2021 IEEE Wireless Communications and Networking Conference (WCNC), Nanjing, China, 29 March–1 April 2021; IEEE: New York, NY, USA, 2021; pp. 1–6.
5. Krunz, M.; Aykin, I.; Sarkar, S.; Akgun, B. Online Reinforcement Learning for Beam Tracking and Rate Adaptation in Millimeter-Wave Systems. *IEEE Trans. Mob. Comput.* **2024**, *23*, 1830–1845. [\[CrossRef\]](#)
6. Sun, G.; Liu, Y.; Chen, Z.; Wang, A.; Zhang, Y.; Tian, D.; Leung, V.C. Energy efficient collaborative beamforming for reducing sidelobe in wireless sensor networks. *IEEE Trans. Mob. Comput.* **2019**, *20*, 965–982. [\[CrossRef\]](#)
7. Eappen, G.; Cosmas, J.; Nilavalan, R.; Thomas, J. Deep learning integrated reinforcement learning for adaptive beamforming in B5G networks. *IET Commun.* **2022**, *16*, 2454–2466. [\[CrossRef\]](#)
8. Arjoune, Y.; Faruque, S. Experience-driven learning-based intelligent hybrid beamforming for massive MIMO mmWave communications. *Phys. Commun.* **2022**, *51*, 101534. [\[CrossRef\]](#)
9. Lin, J.; Zout, Y.; Dong, X.; Gong, S.; Hoang, D.T.; Niyato, D. Deep reinforcement learning for robust beamforming in IRS-assisted wireless communications. In Proceedings of the GLOBECOM 2020—2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020; IEEE: New York, NY, USA, 2020; pp. 1–6.
10. Chen, J.; Feng, W.; Xing, J.; Yang, P.; Sobelman, G.E.; Lin, D.; Li, S. Hybrid beamforming/combining for millimeter wave MIMO: A machine learning approach. *IEEE Trans. Veh. Technol.* **2020**, *69*, 11353–11368. [\[CrossRef\]](#)
11. Khan, M.Q.; Gaber, A.; Schulz, P.; Fettweis, G. Machine Learning for Millimeter Wave and Terahertz Beam Management: A Survey and Open Challenges. *IEEE Access* **2023**, *11*, 11880–11902. [\[CrossRef\]](#)
12. Huang, C.; Alexandropoulos, G.C.; Zappone, A.; Yuen, C.; Debbah, M. Deep learning for UL/DL channel calibration in generic massive MIMO systems. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; IEEE: New York, NY, USA, 2019; pp. 1–6.
13. ElHalawany, B.M.; Hashima, S.; Hatano, K.; Wu, K.; Mohamed, E.M. Leveraging Machine Learning for Millimeter Wave Beamforming in Beyond 5G Networks. *IEEE Syst. J.* **2022**, *16*, 1739–1750. [\[CrossRef\]](#)

14. Liu, S.; Sun, G.; Li, J.; Liang, S.; Wu, Q.; Wang, P.; Niyato, D. UAV-enabled Collaborative Beamforming via Multi-Agent Deep Reinforcement Learning. *arXiv* **2024**, arXiv:2404.07453. [\[CrossRef\]](#)
15. Zhou, X.; Chen, X.; Tong, L.; Wang, Y. Attention-deep reinforcement learning jointly beamforming based on tensor decomposition for RIS-assisted V2X mmWave massive MIMO system. *Complex Intell. Syst.* **2024**, *10*, 145–160. [\[CrossRef\]](#)
16. Sapavath, N.N.; Safavat, S.; Rawat, D.B. On the machine learning-based smart beamforming for wireless virtualization with large-scale MIMO system. *Trans. Emerg. Telecommun. Technol.* **2019**, *30*, e3713. [\[CrossRef\]](#)
17. Kaur, J. Exploiting the Location Information for Adaptive Beamforming in Transport Systems. PhD Thesis, University of Glasgow, Glasgow, UK, 2024.
18. Zhang, T.; Wen, H.; Jiang, Y.; Tang, J. Deep-Reinforcement-Learning-Based IRS for Cooperative Jamming Networks Under Edge Computing. *IEEE Internet Things J.* **2023**, *10*, 8996–9006. [\[CrossRef\]](#)
19. Lu, S.; Shen, X.; Zhang, P.; Wu, Z.; Chen, Y.; Wang, L.; Xie, X. Deep Reinforcement Learning-Based Intelligent Reflecting Surface for Cooperative Jamming Model Design. *IEEE Access* **2023**, *11*, 98764–98775. [\[CrossRef\]](#)
20. Bhuyan, M.; Sarma, K.K.; Misra, D.D.; Guha, K.; Iannacci, J. Adaptive Hybrid Beamforming Codebook Design Using Multi-Agent Reinforcement Learning for Multiuser Multiple-Input–Multiple-Output Systems. *Appl. Sci.* **2024**, *14*, 7109. [\[CrossRef\]](#)
21. Ge, J.; Liang, Y.C.; Zhang, L.; Long, R.; Sun, S. Deep Reinforcement Learning for Distributed Dynamic Coordinated Beamforming in Massive MIMO Cellular Networks. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 4155–4169. [\[CrossRef\]](#)
22. Liang, S.; Fang, Z.; Sun, G.; Liu, Y.; Qu, G.; Jayaprakasam, S.; Zhang, Y. A joint optimization approach for distributed collaborative beamforming in mobile wireless sensor networks. *Ad Hoc Netw.* **2020**, *106*, 102216. [\[CrossRef\]](#)
23. Liu, C.; Helgert, H.J. An improved adaptive beamforming-based machine learning method for positioning in massive mimo systems. *Int. J. Adv. Internet Technol.* **2020**, *6*, 1–12.
24. Elbir, A.M.; Mishra, K.V.; Vorobyov, S.A.; Heath, R.W. Twenty-five years of advances in beamforming: From convex and nonconvex optimization to learning techniques. *IEEE Signal Process. Mag.* **2023**, *40*, 118–131. [\[CrossRef\]](#)
25. Hamza, S.A.; Juretus, K.; Amin, M.G. Sparse Array Design for Optimum Beamforming Using Deep Learning. In *Sparse Arrays for Radar, Sonar, and Communications*; Wiley: Hoboken, NJ, USA, 2024; pp. 215–250.
26. Foundation, P.S. Python Programming Language—Official Website. 2024. Available online: <https://www.python.org/> (accessed on 19 September 2024).
27. Developers, T. TensorFlow—Official Website. 2024. Available online: <https://www.tensorflow.org/> (accessed on 19 September 2024).
28. Anaconda, I. Anaconda—Official Website. 2024. Available online: <https://www.anaconda.com/> (accessed on 19 September 2024).
29. Developers, K. Keras—Official Website. 2024. Available online: <https://keras.io/> (accessed on 19 September 2024).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.