Deep Q-Networks in Adaptive k-space Sampling for Precision MRI

Aliaksei Kaliutau[1]

Imperial College London, London, UK
aliaksei.kaliutau24@imperial.ac.uk

Abstract. This paper presents a novel Deep Reinforcement Learning approach for optimizing k-space vertical line selection in undersampled Magnetic Resonance Imaging (MRI). By training a Deep Q-Network (DQN) agent to directly learn a sampling policy from reconstruction quality feedback (PSNR and SSIM), we demonstrate its ability to discover strategies superior to common heuristics, especially in high uncertainty regimes. Experimental results demonstrate significant quantitative gains, suggesting that actively tailoring sampling patterns to specific image content and the reconstruction process through RL can substantially improve accelerated MRI performance within a fixed acquisition budget. This work highlights the potential of reinforcement learning for precise image reconstruction within the MRI workflow, opening promising avenues for further improvement of clinical MRI technologies.

Keywords: Image Reconstruction · Precision MRI · Deep Reinforcement Learning · Scan Trajectory · Fast MRI.

1 Introduction

Magnetic Resonance Imaging (MRI) is a cornerstone of modern medical diagnostics, providing high-quality soft tissue contrast resolution without ionizing radiation [10]. However, a significant limitation of MRI is a slow data acquisition process, which can lead to discomfort in the patient and motion artifacts, limiting its clinical utility in emergency settings or for dynamic physiological processes. Although there are fast technologies that allow reconstruct low-resolution MRI images relatively quickly (in 1-5 seconds), acquiring sufficient data in the k-space to reconstruct high-resolution, low-noise images often requires significantly longer scan times ranging from minutes to tens of minutes [1]. Consequently, reducing MRI scan times while preserving acceptable diagnostic image quality remains a paramount challenge in the field.

A primary strategy to accelerate MRI is k-space undersampling, when only a fraction of the normally required data points is acquired [7]. This proportionally reduces scan time, but a simple direct reconstruction using methods such as inverse FFT results in aliasing artifacts obscuring anatomical details and potentially compromising diagnosis. The methods capable of mitigating these artifacts often fall into two categories: (i) image-domain approaches, which typically learn mappings from aliased under-sampled images to their fully-sampled counterparts, often using deep learning architectures like U-Nets [6], filters and classic ML approaches, and (ii) k-space domain approaches, which aim to estimate the missing k-space data landscape before image reconstruction, leveraging techniques like Compressed Sensing (CS) or deep learning models operating directly in the frequency domain [2].

Regardless of the approach, the effectiveness of any reconstruction method is fundamentally linked to the information content captured by the acquired k-space samples, since the k-space is the primary and fundamental source of data. This information is determined by the sampling patterns, the specific locations in the k-space that are measured ("sampled"). Traditional accelerated MRI often employs predefined, heuristic sampling patterns, such as variable-density Cartesian or radial trajectories, designed to capture low-frequency information densely (k-space center) while sparsely sampling high frequencies (k-space periphery). Although effective to some extent, these fixed patterns may be suboptimal in maximizing image quality with minimal acquisition time[8].

This limitation motivates a shift towards adaptive or dynamic k-space sampling strategies[3], [4], [8]. The main idea is that by dynamically focusing attention on the most informative k-space regions for a *specific image instance* or reconstruction task, a superior reconstruction quality can be achieved for a given scan time budget, compared to fixed sampling approaches. This paradigm transforms the acquisition process itself into an optimization problem.

In this paper, we investigate the potential of dynamic k-space sampling guided by Reinforcement Learning (RL) and analyze how different sampling masks affect the quality of image reconstruction. Our contributions are twofold. First, we define the sequential selection of k-space scan lines as a sequential decision-making problem. We propose and implement a Deep Q-Network (DQN) agent [5] trained to learn an optimal policy to dynamically select the next line to acquire, based on the current state of the partially acquired k-space. The objective of the RL agent is to maximize image quality metrics, such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [9], within a predefined acquisition budget (e.g., total number of lines acquired, N). Second, we perform comparative analyzes evaluating the impact of different sampling strategies – including fixed heuristic patterns and our proposed RL-driven dynamic approach – on reconstruction quality metrics (PSNR, SSIM) using the whole validation dataset. In that way, we explore the fundamental interplay between sampling strategy and reconstruction quality in the context of accelerated MRI and trustworthy AI in medical imaging.

2 Methods

2.1 Problem Formulation as a Markov Decision Process (MDP)

We focus on a common scenario where k-space data is acquired sequentially line-by-line, specifically acquiring vertical lines in k-space. Although more sophisticated scan line shapes could be employed, we utilize Cartesian undersampling as the most hardware-friendly approach. In this mode, the MRI scanner's gradients switch on and off along one axis, while another axis remains constant for each scan line. This is the standard approach, and most MRI scanners are optimized for it. [10]

Mathematically our problem can be formulated as follows. Given a fixed budget of vertical lines N to acquire (out of possible lines M) and an initial acquisition of central lines K, the problem is to determine the optimal sequence to select the remaining $N - K$ lines to maximize the quality of the final reconstructed image (or to find the shortest sequence which results in satisfactory reconstruction). Image quality is quantified using standard metrics such as PSNR and SSIM[9].

We formulate the problem of optimal vertical k-space line selection as an MDP, defined by the policy $\pi(S, A, P, R, \gamma)$:

- State Space (S): The state s_t at step t represents the current undersampling mask.
- Action Space (A): In state s_t, an action a_t corresponds to selecting the index i of an additional vertical line to add to the mask. The set of available actions $A(s_t)$ consists of all indices i such that $s_t[i] = 0$ (e.g., not taken yet line). Thus, $a_t \in \{i|s_t[i] = 0\}$.
- Transition Probability (P): The state transition is deterministic. If the agent takes action $a_t = i$ in state s_t, the next state s_{t+1} is obtained by setting the i-th element of the mask vector to 1: $s_{t+1}[i] = 1$ and restoring the image from k-space using this undersampling mask.
- Reward Function (R): The reward R_{t+1} is calculated after taking action a_t and transitioning to state s_{t+1}. It reflects the improvement in reconstruction quality achieved by adding the selected line i. Let $I(s_t)$ be the image reconstructed using the mask represented by state s_t, and $Q(I(s_t))$ be its quality (e.g., PSNR or SSIM) compared to the fully sampled ground truth image. The immediate reward is defined as the change in quality: $R_{t+1} = Q(I(s_{t+1})) - Q(I(s_t))$ This encourages the agent to select lines that produce the greatest immediate improvement in the quality of reconstruction, contributing toward the overall goal of maximizing the final quality.
- Discount Factor (γ): A discount factor $\gamma \in (0, 1]$ determines the importance of future rewards relative to immediate rewards.

2.2 Deep Q-Network (DQN) Agent

We employ a DQN agent [5] to learn the optimal action-value function, $Q(s, a)$, which estimates the expected cumulative discounted reward starting from state s, taking action a, and following the optimal policy thereafter. The DQN uses a deep neural network as an approximation for the Q-function.

Network Architecture The input to the Q-network is the state representation s (which is represented by the tensor of shape (320,), where 320 is the resolution of image along one axis). The output layer provides Q-value estimates for each possible action (selecting any of the $M = 320$ line indices). During action selection and Q-value updates, outputs corresponding to invalid actions (e.g., lines already present in the mask s) are masked or ignored.

Q-Value Update Rule The network's weights are updated iteratively using experiences sampled from a replay buffer. The core of the learning process relies on minimizing the difference between the predicted Q-value $Q(s,a)$ and a target value derived from the Bellman equation. In DQN, the target value for an experience tuple (s,a,R,s') is often calculated as [5]:

$$Q(s,a) = R + \gamma * \max_{a'} Q_{target}(s',a'),$$

where:

- $Q(s,a)$: The current Q-value estimate from the main network for taking action a (selecting a specific line) in state s (current mask).
- R: The immediate reward obtained after taking action a, transitioning to state s', and performing the reconstruction and quality evaluation.
- γ: The discount factor, to balance immediate versus future rewards.
- s': The state (mask) resulting from adding line a to mask s.
- $max_{a'} Q_{target}(s',a')$: This term represents the maximum expected future cumulative reward achievable from the next state s', according to the target network (Q_{target}). It estimates the value of the best action a' (e.g., the best next line to pick) that can be taken from state s'.1

The update rule aims to make the Q-value $Q(s,a)$ accurately reflect the true value for selecting line a when the current sampling mask is s. As a result, the agent learns which sequences of line selections lead to the highest cumulative rewards, corresponding to the best final image quality within the budget N. We expect this approach to lead to the discovery of scan patterns that will improve the quality of MRI reconstruction.

Training We use standard DQN techniques such as experience replay, and ϵ-greedy exploration[5]. Exploration (parameterized by ϵ) allows the agent to occasionally choose random lines, discovering potentially valuable actions it might otherwise miss.

2.3 Limitations

In our work, we did not aim to create a scalable RL agent capable of learning from large datasets. Instead, we trained our agent on a few carefully selected samples that were manually evaluated and added. However, the sampling masks generated by the agent were validated against the entire test dataset (500 images). Consequently, the agent is likely suboptimal, and our results likely represent a lower bound on potential performance. (See also the "Further Work" section).

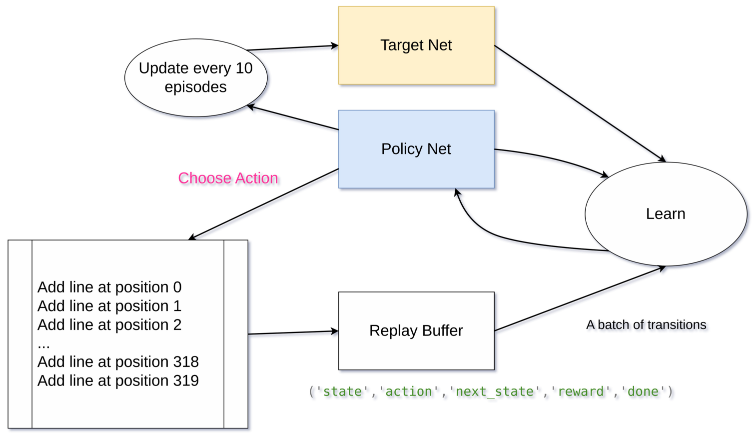

Fig. 1. Actions are selected randomly or based on a policy, and the resulting *state transitions* are stored in Replay Buffer. At each iteration, an optimization step is performed, using a random batch from Replay Buffer to train the policy. The older Target Network is used to compute expected Q-values during optimization. A soft update is applied to the target network's weights every 10 episodes.

3 Experiments

3.1 Experimental Setup

Dataset Experiments were conducted on the FastMRI dataset [11]. All images were pre-processed using normalizations to [0, 1]. We used 1-5 images for training and 500 validation images for testing.

Initial sampling mask/state The initial state s_0 always included the central K vertical lines of the k-space. We set $K = 4$. The total number of vertical lines to be acquired was set to $N = 20$ and $N = 32$ for two separate experiments.

DQN Agent Network: The DQN network is a convolutional NN designed to estimate Q-values for reinforcement learning. It takes a single-channel input and processes it through three 1D convolutional layers (with kernel sizes 4, 2, and 2, respectively) followed by ReLU activations. The output of the final convolutional layer is flattened and then passed through two fully connected layers, with the final layer outputting Q-values for each of the 320 possible actions.

Hyperparameters: Learning rate α = 1e-3, discount factor $\gamma = 0.95$, replay buffer size = 1000, batch size = 32, target network update frequency = 10, ϵ-greedy exploration schedule was set to linearly decaying from 1.0 to 0.05 over training steps.

Evaluation Metrics: The primary metrics were PSNR and SSIM, calculated between the reconstructed image using the N selected lines and the fully sampled ground truth image reconstructed from $M = 320$ scan lines.

3.2 Baselines

We compared the performance of our DQN agent against the following baseline selection strategies:

Random: After acquiring the initial K central lines, the subsequent $N - K$ lines were chosen uniformly at random without replacement of the remaining available lines.

Equispaced: Lines were selected such that the final set of N lines (including the initial K) were approximately equally spaced across the entire k-space width.

Low-Frequency Biased: Lines were selected sequentially outwards from the initial K central block, adding adjacent lines until the budget N was reached.

The baseline masks were generated using recipe and code from FastMRI project, in particular classes `RandomMaskFunc` and `EquispacedMaskFractionFunc` [11]

3.3 Results

We evaluated the performance by averaging the final PSNR and SSIM values over the test dataset for each method.

We observed that the DQN approach achieved an average PSNR of 22.25 dB and SSIM of 0.4603 (total $N = 20$ lines) and PSNR of 24.09 dB and SSIM of 0.5275 ($N = 32$ lines), significantly outperforming Random, Equispaced, and Low-Frequency Biased masks (see Table. 1, 2).

Table 1. Reconstruction quality for masks with the budget of N=20 scan lines.

Mask name	PSNR	σ	SSIM	σ
rl_4_20_train (ours)	**22.25**	**1.711**	**0.4603**	**0.0962**
fMRI_Ran_AF16_CF0.02_PE320	19.87	1.741	0.4140	0.0903
fMRI_Reg_AF16_CF0.02_PE320	19.74	1.737	0.4116	0.0907

Table 2. Reconstruction quality for masks with the budget of N=32 scan lines.

Mask name	PSNR	σ	SSIM	σ
rl_4_32_train (ours)	**24.09**	**1.820**	**0.5275**	**0.0895**
fMRI_Ran_AF8_CF0.04_PE320	22.55	1.814	0.4824	0.0917
fMRI_Reg_AF8_CF0.04_PE320	22.62	1.805	0.4986	0.0876

Figure 2 illustrates that the DQN agent learned a strategy that yielded consistent improvements in reconstruction quality with each selected line, generally achieving better quality than baselines within scan line budget of N lines.

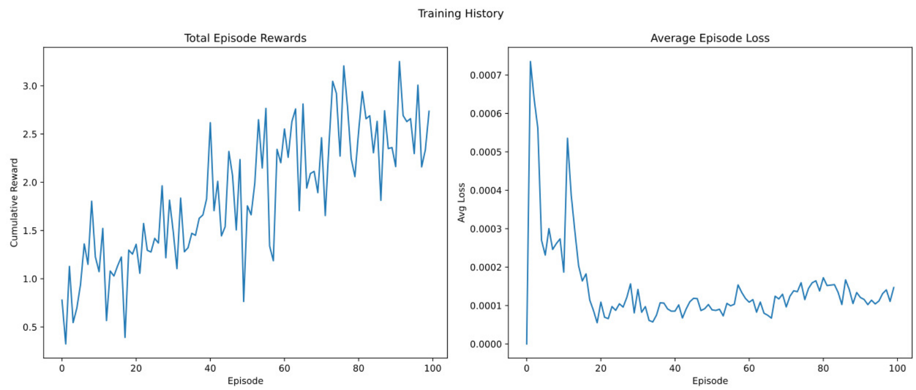

Fig. 2. The DQN agent learns where to add a line to the sampling mask to achieve the greatest possible improvement in image quality.

4 Discussion

The experimental results demonstrate the effectiveness of the proposed DQN-based approach for optimizing k-space vertical line selection in undersampled MRI. By learning a policy directly from the reconstruction quality feedback, the RL agent discovered line selection strategies superior to common heuristics. The significant quantitative gains in PSNR and SSIM scores suggest that actively tailoring the sampling pattern based on the specific image content and reconstruction process, guided by RL, can lead to substantially improved accelerated MRI performance within a fixed acquisition budget. The learned policy likely captures complex dependencies between k-space line locations and their contribution to reducing reconstruction artifacts.

Analyzing the RL agent's heatmap of selection patterns, we found a consistent strategy: the agent tended to modestly expand the sampling density in the central k-space region, crucial for overall image contrast and structure, while strategically acquiring sparser lines in higher-frequency peripheral regions to capture finer details and edges. This adaptive balancing act highlights the agent's ability to move beyond simple heuristics.

Interestingly, the learned policies consistently deviated from periodic, uniform undersampling patterns. The agent never converged to such a regular pattern in our experiments, strongly suggesting that for the given reconstruction framework and input data, the optimal sampling mask is inherently non-uniform.

In summary, our work emphasizes the potential of reinforcement learning as a powerful tool for optimization tasks within medical imaging. It is worth to note, that the k-space line selection is a small task in the bigger workflow, and harnessing RL for optimization opens promising avenues for further improvement of clinical MRI technologies.

Our study demonstrates the promise of RL for MRI sampling optimization, but several avenues for future research remain. In particular, we plan to focus on two key areas.

- Scalability and Generalization. The current approach should be updated to train agent on large diverse datasets, rather than using a few highly selected images for that.
- Acquisition Budget schemes. We hypothesize that the advantages of RL-based optimization will be particularly pronounced in ultra-low budget regimes (i.e., in the mode of very high MRI acceleration factors), where the strategic selection of each k-space line becomes critically important for preserving essential diagnostic information. Quantifying performance gains in these challenging scenarios would be highly valuable and should be a topic to further research the Trustworthy AI aspect in this project.

References

1. Avey, G.D., Belden, D.S., Zea, R.D., Yu, J.P.J.: Factors predicting the time-length variability of identically protocoled MRI exams. J. Magn. Reson. Imaging **49**(7), e265–e270 (Jun 2019)
2. Jin, K.H., McCann, M.T., Froustey, E., Unser, M.: Deep convolutional neural network for inverse problems in imaging. IEEE Transactions on Image Processing **26**(9), 4509–4522 (Sep 2017). https://doi.org/10.1109/tip.2017.2713099
3. Liu, Y., Pang, Y., Jin, R., Hou, Y., Li, X.: Reinforcement Learning and Transformer for Fast Magnetic Resonance Imaging Scan. IEEE Transactions on Emerging Topics in Computational Intelligence **8**(3), 2310–2323 (2024). https://doi.org/10.1109/TETCI.2024.3358180
4. Liu, Y., Pang, Y., Sun, X., Hou, Y., Yang, Z., Wang, Z.: Dual states based reinforcement learning for fast MR scan and image reconstruction. Neurocomput. **568**(C) (Feb 2024). https://doi.org/10.1016/j.neucom.2023.127067
5. Mnih, V., et al.: Human-level control through deep reinforcement learning. Nature **518**(7540), 529–533 (Feb 2015). https://doi.org/10.1038/nature14236
6. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation (2015), https://arxiv.org/abs/1505.04597
7. Seeger, M., Nickisch, H., Pohmann, R., Scholkopf, B.: B.: Optimization of k-space trajectories for compressed sensing by bayesian experimental design. Magnetic resonance in medicine **63**, 116–26 (11 2009). https://doi.org/10.1002/mrm.22180
8. Sherry, F., Benning, M., De Los Reyes, J.C., Graves, M.J., Maierhofer, G., Williams, G., Schonlieb, C.B., Ehrhardt, M.J.: Learning the sampling pattern for MRI. IEEE Trans. Med. Imaging **39**(12), 4310–4321 (Dec 2020)
9. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004). https://doi.org/10.1109/TIP.2003.819861
10. Westbrook, C., Talbot, J.: MRI in Practice. Wiley Publishing, 5th edn. (2018)
11. Zbontar, J., et al.: fastMRI: An Open Dataset and Benchmarks for Accelerated MRI (2019), https://arxiv.org/abs/1811.08839