RL proposal: Agent finds the tumor in 3D MRI

Su-Min Roh
20234071
rohsumin325@gm.gist.ac.kr

Han-Jun Choi
20241049
gkswns3708@gm.gist.ac.kr

Gumwoo Bae
20241138
gunwoof1234@gm.gist.ac.kr

I. Introduction

Reinforcement learning (RL) is a branch of machine learning that focuses on learning through interactions to maximize rewards. Its application to accurate diagnosis and efficient treatment plans using medical images, especially 2D and 3D MRI images for brain tumor localization, has shown considerable promise [3], [4].

Brain tumors affect their neurological function depending on their size and location. Accurate determination of the tumor's boundaries is critical for effective treatment. Our study detects brain tumors from 3D MRI images by applying RL to help us understand the complex structure of brain tumors in more detail.

II. Approach

A. Dataset

The BraTS 2020 dataset includes pre-surgical multimodal MRI scans from 19 institutions, with four modalities (T1, T1Gd, T2, T2-FLAIR) from about 400 patients [1]. Annotations, by clinicians, cover several tumor regions with essential imaging preprocessing like co-registration and skull-stripping already applied.

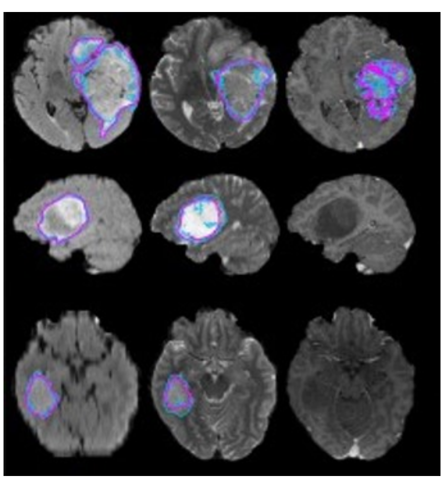

Fig. 1. The examples of the BraTS data, with tumor regions as inferred from the annotation of experts (blue lines) and consensus segmentation (purple liens).

B. Method

The whole architecture of this project is shown in Fig 2.

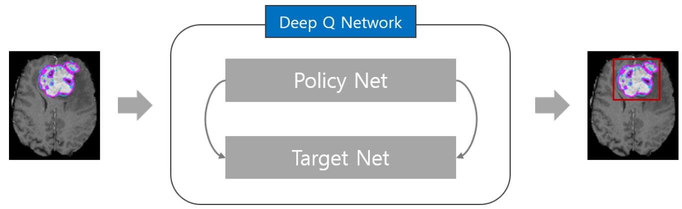

Fig. 2. The whole architecture that will be used in our project.

1) State Space: In this project, the state space is defined by the position and order of the bounding box in the 3D MRI image, which depends on each action taken by the agent. This dynamic adjustment allows the agent to continuously update the focus within the image while interacting with the environment.

2) Action Space: The action space contains a set of movements (up, down, left, right, up, down, down) that allow the agent to manipulate the bounding boxes to align them more accurately to the tumor. Each action will be designed to improve detection accuracy by adjusting the position of the bounding boxes according to the agent's current state.

3) Reward: The compensation is calculated based on the proximity of the bounding box to the actual tumor location and is quantified by the bounding cross (IoU) score. The IoU measures the overlap between the bounding box (bb) and the actual tumor location (gt) coordinated by the agent. The compensation function then evaluates the changes in the IoU, allowing the agent to find a more accurate location: $R_a(s, s')$ is the reward for transitioning from one state to another, and bb' is the newly adjusted bounding box. Improved IoU scores increase the reward, providing a positive reinforcement to behaviors that improve agents' tumor detection accuracy.

4) Transition Matrix: The Markov Decision Process (MDP) used in this project is a key element of the transition matrix, which details the transition probability based on the current state and the actions taken. We summarize the agent's decision-making process by modeling how the positioning of the bounding box affects the transition to the next state.

5) Action evaluation: The action evaluation is performed using the ϵ-greedy algorithm, which starts with random action selection and transitions to increasingly strategic actions based on the search speed (ϵ-greedy = 0.2). This approach allows agents to balance exploring new strategies with exploiting known successful maneuvers, facilitating optimal learning paths.

6) Policy estimation: Policy estimation with Q-learning, a model-free approach, involves direct interactions with the environment where only state and related rewards guide action decisions. Supporting this, the Deep Q-Network (DQN) model includes two main components:

- **PolicyNet**: Learns the action-value function $Q(s, a)$ after each episode.
- **TargetNet**: Periodically synchronizes with PolicyNet to stabilize the learning.

After the follow-up, the agent updates the Q value using the received rewards and the maximum Q value of the subsequent states according to the Bellman equation. The learning rate (α) and discount rate (γ) parameters dominate this update process.

7) Model evaluation: Model evaluation measures the effectiveness of policies in real-world scenarios, measuring metrics such as success rate, IoU score, and average compensation per episode. These metrics reflect the agent's accuracy to tumor location, the accuracy of bounding box alignment with tumors, and the overall improvement of agent performance through the learning process.

III. Conclusion

This project adopted the ideas of two previous studies to enhance tumor localization with 3D MRI images using reinforcement learning [2], [4]. The first study, in which the results are shown in Fig 3, could lose some information because it used 2D slices from 3D MRI images to detect tumors [4]. The focus of the second study was to locate the hippocampus for Alzheimer's prediction [2], the result is shown in Fig 4. In this project, we transformed hippocampus localization into tumor finding. We aim to leverage the insights of these studies to improve the accuracy and effectiveness of medical images for tumor diagnosis and treatment planning, laying the foundation for future developments in this critical area.

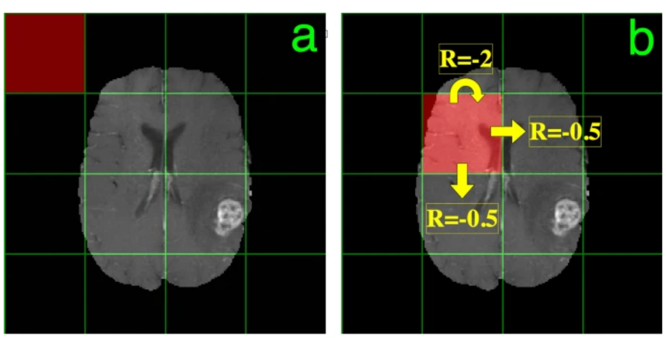

Fig. 3. The result of the first related work, from 'Reinforcement learning using Deep networks and learning accurately localizes brain tumors on MRI with very small training sets' [4].

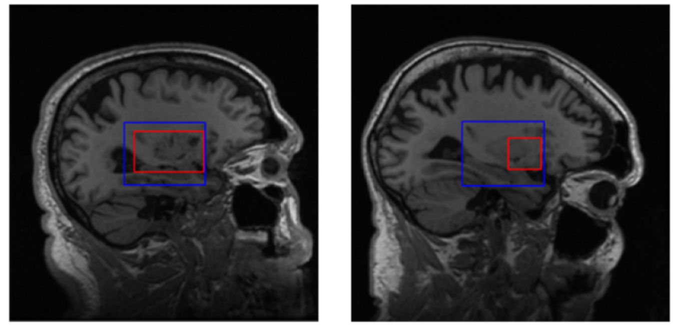

Fig. 4. The process of the second related work, from 'Reinforcement-Learning-Based Localization of Hippocampus for Alzheimer's Disease Detection' [2].

References

[1] Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al.: The multimodal brain tumor image segmentation benchmark (brats). IEEE transactions on medical imaging **34**(10), 1993–2024 (2014)

[2] Raj, A., Mirzaei, G.: Reinforcement-learning-based localization of hippocampus for alzheimer's disease detection. Diagnostics **13**(21), 3292 (2023)

[3] Stember, J., Shalu, H.: Deep reinforcement learning to detect brain lesions on mri: a proof-of-concept application of reinforcement learning to medical images. arXiv preprint arXiv:2008.02708 (2020)

[4] Stember, J.N., Shalu, H.: Reinforcement learning using deep q networks and q learning accurately localizes brain tumors on mri with very small training sets. BMC Medical Imaging **22**(1), 224 (2022)