Masoud Nateghi

BMI 539: Introduction to Reinforcement Learning

Paper 1: Ghesu FC, Georgescu B, Mansi T, Neumann D, Hornegger J, Comaniciu D. An artificial agent for anatomical landmark detection in medical images. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016, Cham. 2016.

Problem description: Landmark detection in biomedical image analysis refers to identifying specific anatomical points, or biological coordinates, that can be consistently and precisely located across images obtained from various imaging modalities, such as CT, ultrasound, and MRI. Accurate detection of these landmarks is crucial for subsequent medical image analysis tasks.

State/action/reward definition: In an image I, $\overrightarrow{P_{GT}}$ denoted the location of anatomical landmark, and $\overrightarrow{p_t}$ denoted the location at the current time. State space S was the collection of all possible states$s_t = I(\overrightarrow{p_t})$. Action space A was the collection of all possible actions by which the agent can move to the adjacent pixels. Reward space R was defined as $||\overrightarrow{p_t} - \overrightarrow{p_{GT}}||_2^2 - ||\overrightarrow{p_{t+1}} - \overrightarrow{p_{GT}}||_2^2$ which impelled the agent to move closer to the target anatomical landmark.

RL algorithm to solve the problem: Deep Q-learning

Model based or model free: Model free

Paper 2: Maicas G, Carneiro G, Bradley AP, Nascimento JC, Reid I, Deep reinforcement learning for active breast lesion detection from DCE-MRI. International conference on medical image computing and computer-assisted intervention. 2017.

Problem description: Lesion detection (object detection) involves identifying the class labels and locations of target objects within images or videos, often referred to as object extraction. This process is a key task in medical image analysis.

State/action/reward definition: The states were defined by the current bounding box volumes in the 3D dynamic contrast-enhanced MR images. The action set included nine possible actions, allowing the bounding box to move forward or backward along the x-, y-, and z-axes, resize (scale up or down), or trigger the terminal state. A positive reward was given when the agent moved toward the lesion, and a negative reward when it moved away. If the agent remained within the lesion area, it received a large positive reward, whereas staying outside the lesion area resulted in a large negative reward.

RL algorithm to solve the problem: Deep Q-network (DQN)

Model based or model free: Model free

Paper 3: Navarro F, Sekuboyina A, Waldmannstetter D, Peeken JC, Combs SE, Menze BH. Deep reinforcement learning for organ localization in CT. Medical Imaging with Deep Learning. 2020.

Problem description: Organ/anatomical structure detection to locate various organs in 3D CT scans.

State/action/reward definition: The state was defined by the voxel values within the current 3D bounding box. The action set consisted of eleven actions: six for translation, two for zooming, and three for scaling, ensuring that the bounding box could move to any region of the 3D scan. The agent received a reward when an action improved the intersection over union (IoU) score.

RL algorithm to solve the problem: Deep Q-learning

Model based or model free: Model fre

Paper4: Yang H, Shan C, Kolen AF, Deep Q-Network-Driven Catheter Segmentation in 3D US by Hybrid Constrained Semi-Supervised Learning and Dual-UNet. International Conference on Medical Image Computing and Computer-Assisted Intervention. 2020.

Problem description: Image segmentation. Most supervised learning-based catheter segmentation methods rely on a large amount of well-annotated data. This paper used a DQN agent to identify the coarse location of the catheter, followed by patch-based segmentation using Dual-UNet.

State/action/reward definition: The states were defined as 3D observation patches, and the agent could update the states by moving the patch's center point (x, y, z) along the x-, y-, and z-axes within the observation space. Similar to landmark detection tasks, the agent received a negative reward if the patch moved away from the target, a positive reward if it moved toward the target, and no reward if it remained stationary

RL algorithm to solve the problem: Deep Q-learning

Model based or model free: Model free

Paper 5: Ye J, Xue Y, Long LR, et al, Synthetic sample selection via reinforcement learning. International Conference on Medical Image Computing and Computer-Assisted Intervention. 2020.

Problem description: Medical image synthesis involves generating artificial images that contain simulated lesions or anatomical structures across one or multiple imaging modalities. Rather than solely focusing on creating synthetic images, reinforcement learning (RL) can also be applied to assess and select the most suitable synthetic images

State/action/reward definition: The output of the transformer model was a binary action determining whether a synthetic image should be selected or discarded. The reward system was designed to encourage improved accuracy in the subsequent classification task. No specific states were defined in this process.

RL algorithm to solve the problem: Proximal Policy Optimization (PPO)

Model based or model free: Model free