|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2017-2020 | | FG-AI4H-D-018 | |
| **ITU-T Focus Group on AI for Health** | |
| **Original: English** | |
| **WG(s):** | | N/A | Shanghai, 02-05 April 2019 | |
| **DOCUMENT** | | | | |
| **Source:** | | BioMind | | |
| **Title:** | | Automated Generation of Radiotherapy Treatment Plans using Reinforcement Learning | | |
| **Purpose:** | | Discussion | | |
| **Contact:** | | Hu Yue BioMind China | | +8613810094506 huyue@biomind.ai |
| **Contact:** | | Joe Wu BioMind China | | +86 18600942462 joe.wu@biomind.ai |

|  |  |
| --- | --- |
| **Abstract:** | Besides surgery and chemotherapy, radiotherapy is one of the techniques frequently used in the fight against cancer. Over the years, advanced techniques such as intensity modulated radiotherapy (IMRT) and volumetric modulated arc therapy (VMAT) have been developed for radiation delivery procedure. Regardless of the type of delivery mode, a radiotherapy treatment plan needs to be generated first. We use the approach of reinforcement learning to automate the treatment plan generation process to achieve auto-planning better than most dosimetrists. |

**Overview**

Monaco, a treatment planning system developed by Elekta, was used as our research environment. The planning activity is based on inverse planning strategy, which requires a user defined identification of the cost functions and their respective goals in order to find the best configuration of beam intensities. Monaco offers a set of dose-based and biological cost functions, which consists of target penalty, target EUD, serial model, parallel model, quadratic under/over dose, under/over dose DVH, maximum dose and conformality. The final plan quality depends on the quality of the selection for the cost functions and their goals made by the dosimetrist. Hence, tuning of the parameters of each cost functions will be repeated until a satisfactory plan is achieved.

In this project, we propose to use a reinforcement learning method. By defining an appropriate scoring function to evaluate the performance of each plan, we provide a guidance to the reinforcement learning agent on what is the final goal to achieve since it is trained to maximize the returns received. In other words, through a trial-and-error process, the agent is able to learn how good or bad the current plan is and what are the good actions to take in each situation in order to move towards an optimal plan. These actions include tuning of the parameters for each cost functions and adding and/or removing of cost functions. The agent will then stop the iterative process once a near optimal plan has been produced.

# Impact

The radiotherapy treatment plan generation process is a tedious process as it requires a lot of human interaction and experience. A dosimetrist has the task to generate a plan which achieves a list of clinical requirements by selecting cost functions and their goals based on previous experience. The challenge is to find the best balance between ensuring all tumor cells receive a lethal dose of radiation and at the same time having tolerable effect on the other normal tissues. This is a trial-and-error process that is time consuming with no guarantee about the optimality of the plan as it is affected greatly by the dosimetrist’s experience and skills. As the incidence of cancer continues to rise worldwide, the demand for radiation therapy services is expected to increase, naturally requiring more medical professionals in this area. The objective of this project is to generate high quality plans with minimal human interaction with the treatment planning system, and hence solve the problem of inefficiency and shortage of experienced dosimetrists. Furthermore, the quality of the plans generated would be more consistent as it is no longer dependent on the level of experience of the dosimetrist.

# Existing Work

There is no existing solution on automated generation of radiotherapy treatment planning using the approach of reinforcement learning. On the other hand, there are some other work done on this problem which uses non-reinforcement learning methods. Recently, Babier et al.[[1]](#footnote-1) proposed a solution which combines knowledge-based planning (KBP) predictions with an inverse optimization (IO) pipeline. In short, the KBP approaches first predict achievable dose volume histograms, which will then be inputted into an IO pipeline that generated treatment plans via an intermediate step that estimated objective function weights for an inverse planning model. Besides that, Philips Pinnacle has also developed the AutoPlanning software module[[2]](#footnote-2). It uses the iterative approach of progressive optimization as it learns by capturing the steps that a skilled human operator would take. In the algorithm, regional optimization based on region of interest is implemented. In other words, to improve the plan in each iterative step, optimization is emphasized on specific region of interests using relatively high importance factors on small region of interests with the aim to create high-dose gradients between target volumes and organs at risk. This method is based on automatically generated cold and hot regions in the plan.

# Feasibility

Reinforcement learning was inspired by behaviorist psychology with the idea of building a learning system that wants something and can adapt its behavior to maximize a signal it receives from the environment. Unlike supervised learning, there is no supervisor in reinforcement learning, which means there is no guidance on what is the best action to take. Instead, only a reward signal is provided.

It is often that examples of correct and representative behavior of all the available situation are not obtainable in interactive problems, rendering it impractical to implement supervised learning. Instead of being told which actions to take in each situation, the learner, also known as the agent, in reinforcement learning discovers the actions through experience, which leads to the trial-and-error search characteristic. Moreover, the focus of the learner is to achieve the best return in the long run. These characteristics of reinforcement learning make it a perfect fit as a solution to our problem statement.

In recent years, reinforcement learning has gradually become one of the active research areas in the field of machine learning. Hence, many state-of-the-art methods with exceptional performance have been developed, increasing the credibility of using reinforcement learning for our research.

# Data Availability

The datasets available for the project are obtained from the radiotherapy department of the hospital. These data are clinical records of previously treated patients which consist of their CT scans. Besides that, the corresponding target volume and organs at risk delineation for each patient annotated by the physicians are included and it is a required input to Monaco at the start of each treatment plan generation. For each data, their respective clinical requirements are provided by the doctors. These requirements serve as an evaluation of the performance of the automatically generated plan and the goal for the reinforcement learning agent is to achieve them. In other words, having more criteria fulfilled indicates a better quality plan. No other annotations are needed since reinforcement learning agent learns from exploring the possible actions without reference to any explicit labels.

# Benchmarking

For each test case, participants should manually interact with Monaco to generate a treatment plan without prior knowledge of the corresponding final plan suggested by our method or the clinically delivered plan. The performance of each plan will first be evaluated based on the number of clinical requirements satisfied. If the both the manually and automatically generated plan have similar performance, a comparison between the dose distribution calculated based these two plans will be conducted. A plan is considered to be of higher quality if it has a better dose coverage on the target volumes, dose conformality and sparing of the organs at risk.

# Organizer Details

Having an automated treatment planning system can help to greatly reduce the time taken by radiotherapist to create a treatment plan and help inexperience dosimetrist to derive a plan that matches the highly experienced dosimetrist. This the first time that artificial intelligence is applied to radiotherapy treatment planning.

|  |  |
| --- | --- |
| **Organization Name** | BioMind |
| **Contact Name** | Hu Yue |
| **Contact Email Address** | huyue@biomind.ai |
| **Contact Phone Number** | +86 13810094506 |
| **Project Title** | AUTOMATED GENERATION OF RADIOTHERAPY TREATMENT PLANS USING REINFORCEMENT LEARNING |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Babier, Aaron, et al. “Knowledge-based automated planning for oropharyngeal cancer.” *Medical Physics* 45.7(2018) [↑](#footnote-ref-1)
2. Xhaferllari, Ilma, et al. “Automated IMRT planning with regional optimization using planning scripts.” *Journal of Applied Clinical Medical Physics* 14.1(2013):176-191 [↑](#footnote-ref-2)