Multi-Agent Reinforcement Learning Meets Leaf Sequencing in Radiotherapy

Riqiang Gao, Florin C. Ghesu, Simon Arberet, Shahab Basiri, Esa Kuusela, Martin Kraus, Dorin Comaniciu, Ali Kamen

Siemens Healthineers

1. Background

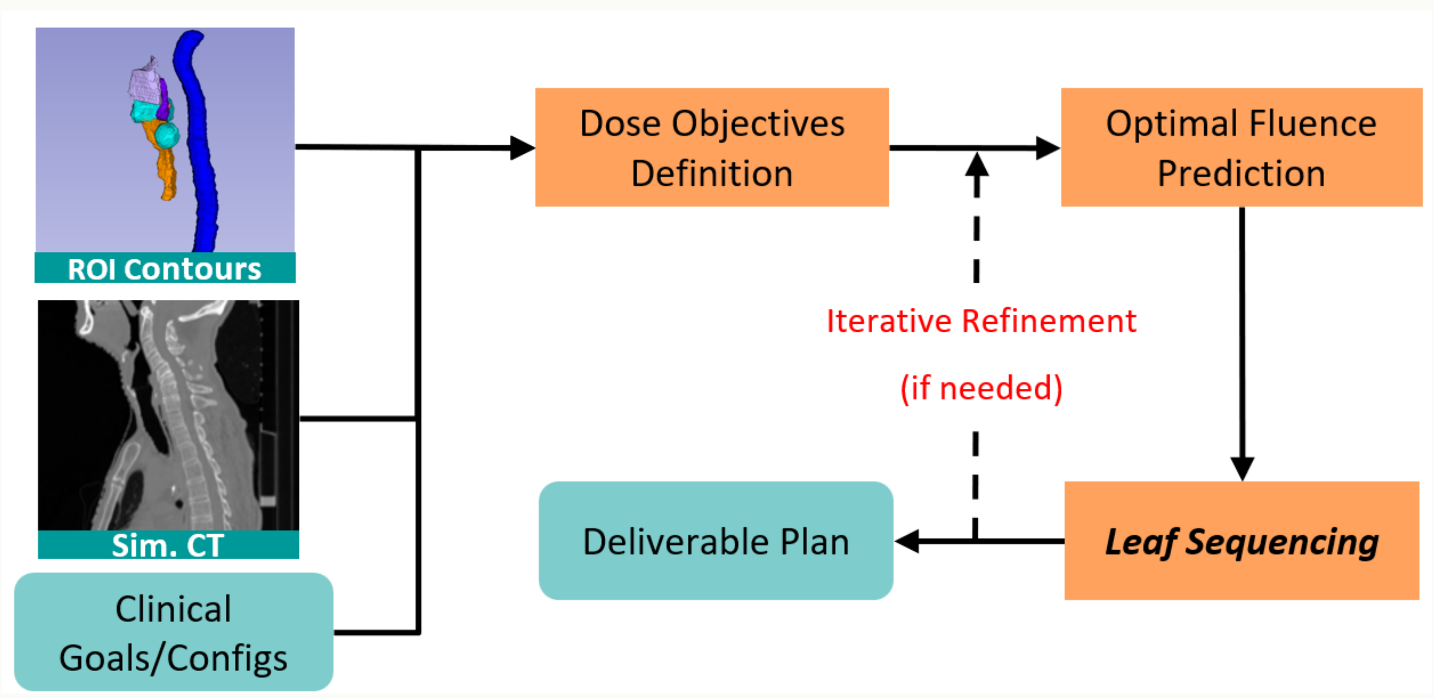

Fig. 1. Illustration of a typical RTP process. Three common components are shown in the orange boxes. We focus on *leaf sequencing* in this work. The term "optimization" in this paper refers to a series of methods that are not machine learning.

2. Motivation

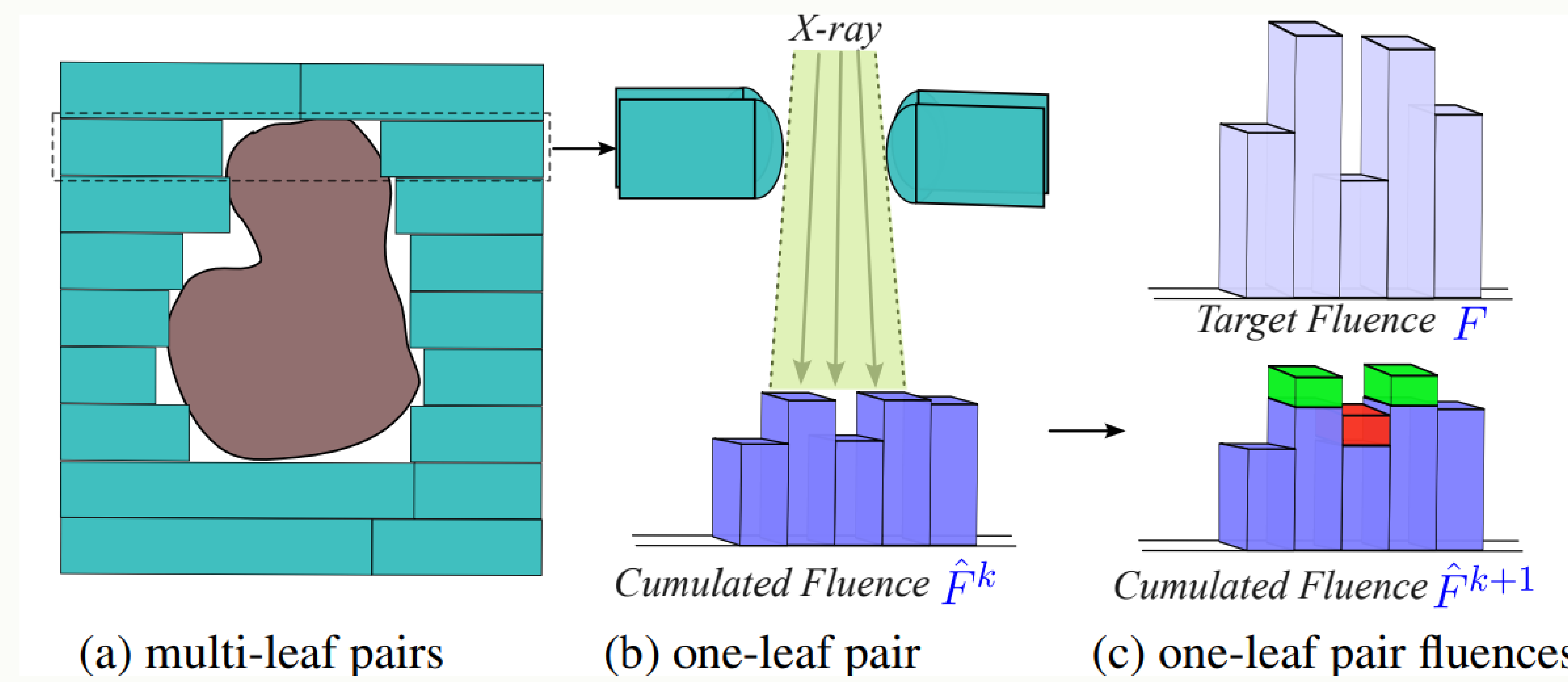

Fig. 2. (a) shows multi-leaf pairs in 2D representing MLC and PTV projections. (b) provides a 3D view of a leaf pair and its connection to cumulated fluences. (c) illustrates motivations of Reward 1 (green) and Reward 2 (red

3. Methodology: Reinforced Leaf Sequencer (RLS)

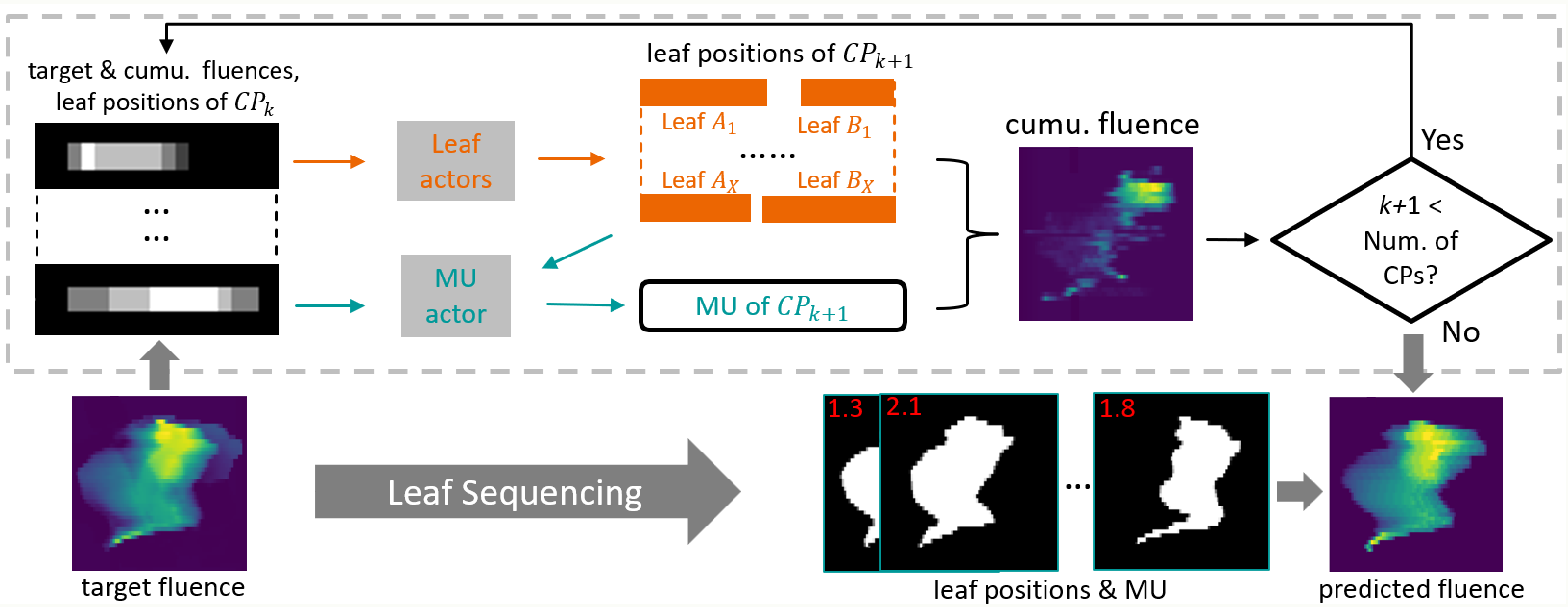

Fig. 3. Illustration of the proposed RLS. The upper shows the methodology and the lower shows the input/output of RLS. The target fluence is splitted into X rows, each row is related to one leaf-pair and one leaf actor. x-th leaf actor predicts the positions of Leaf A_x and B_x. All rows in k-th CP shares the same monitor unit, which is predicted by MU actor after all leaf positions are obtained.

4. Experiment Results

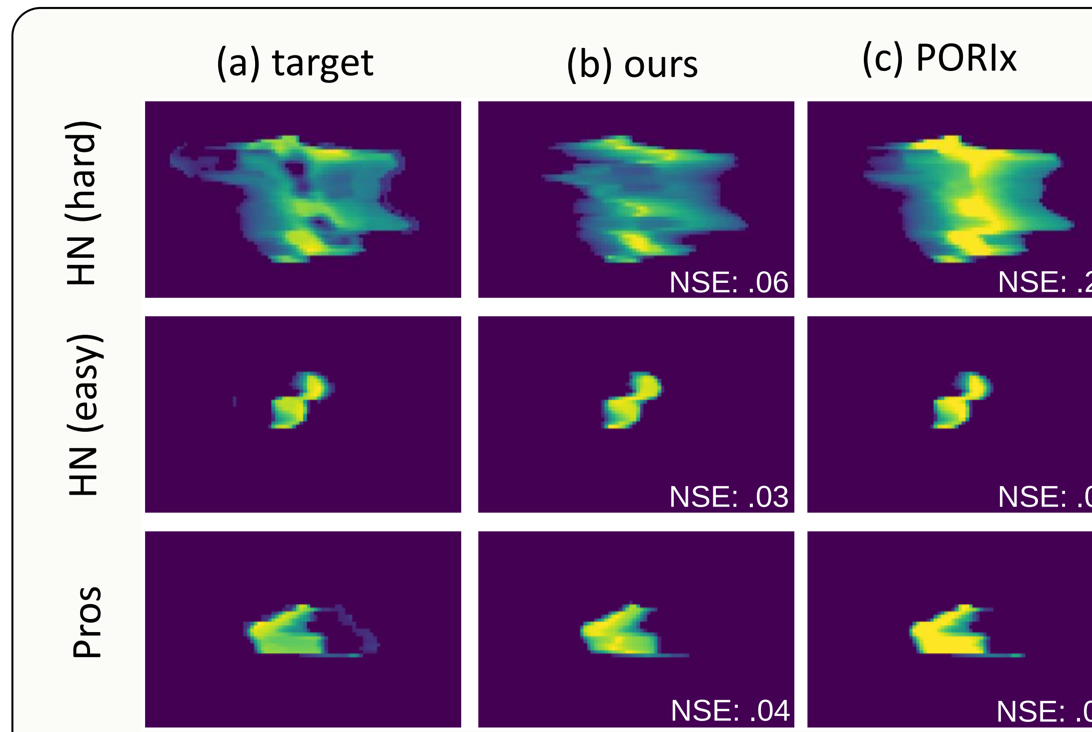

Fig. 4. Predicted fluence vs. target fluence. The reconstruction error is shown right bottom of the predicted fluences. First row: hard case from HN; second row: easy case from HN; third row: prostate case

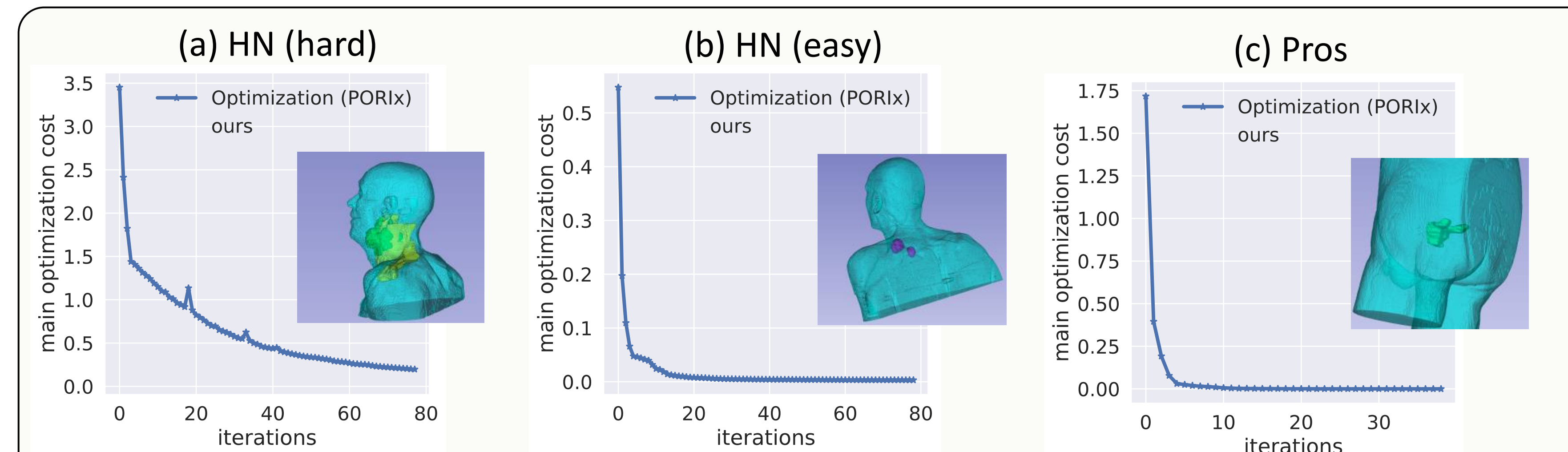

Fig. 5. Typical cases in different scenarios. the PTV contours within body mask are shown along with the main optimization cost vs. the number of iterations. (a): large PTVs (PTV 54 and PTV 60) in HN, (b): small PTVs in HN (PTV 66), (c): PTV in prostate. The RLS brings clear improvements for hard cases (e.g., those with large PTVs).

5. Discussion

- To the best of our knowledge, the proposed RLS is the first MARL-based leaf sequencer for RTP. Limitations and future works have been discussed.
- Excitement surrounds the potential of deep learning to partially or fully replace conventional optimization practical RT in the future.

https://proceedings.mlr.press/v235/gao24g.html