

Online tumor motion forecasting using Vision Transformers

Gauthier Rotsart de Hertaing¹, Damien Dasnoy¹, Guillaume Janssens^{1,2} and Benoît Macq¹

¹Université Catholique de Louvain, Institute of Information and Communication Technologies (ICTEAM), Louvain-La-Neuve, 1348, Belgium

²Ion Beam Applications SA, Louvain-La-Neuve, 1348, Belgium

In radiotherapy, forecasting the motion of a lung tumor is a fundamental task for dose delivery. Indeed, anticipating the motion saves time in the decision-making process for real-time applications and could lead to a better homogeneous dose delivery and more healthy tissues spared. Our method is trained and tested on fluoroscopic sequences generated from 4D-Computed Tomography (4DCT) scans which were acquired one week apart. Due to anatomical changes that can happen during the course of the treatment, training and testing on a different 4DCT is challenging. In that way, we simulate the inter fraction variability that can happen during a treatment over several weeks.

Our method consists of training a Vision Transformer network on fluoroscopy images of the first week of the treatment. Those are cropped around the tumor. Then, the network performance is evaluated on the second week of the treatment. Data augmentation using the software OpenTPS is done and the impact of the number of inter/intra-fractions needed is studied.

This study is realised on fourteen patients with various breathing amplitude (up to 5mm in sagittal, 3mm in axial and 11mm in coronal directions) and different tumor sizes and treatment responses. It yields to an average prediction error of 1.30 mm in the three directions at an horizon of 1s. As future steps, we plan to evaluate the robustness of the Vision Transformer at three to four different timepoints and on breathing amplitude up to 40 mm.

Key-words: Vision Transformer, transformers network, 4DCT, fluoroscopy imaging, radiotherapy, respiratory motion