論 文 の 内 容 の 要 旨

論文題目　　Tumor position prediction using dynamically trained recurrent neural networks for latency compensation in lung cancer radiotherapy

（肺がん放射線治療における遅延補正のための回帰型ニューラルネットワークの適応的学習による腫瘍位置の予測）

氏　　名　　ポル　ミッシェル

Lung tumors follow the respiratory motion and are thus difficult to target accurately during the radiotherapy treatment. Not taking into account the latency of tracking and irradiation systems may lead to uncertainty in the tumor location and high irradiation of healthy tissue surrounding the tumor. This may in turn result in side effects such as radiation pneumonitis or pulmonary fibrosis. Therefore, prediction of the respiratory motion is necessary to deliver radiation to the target precisely. Research concerning respiratory motion prediction for radiation therapy has recently focused on recurrent neural networks (RNN) as their internal feedback loop acts as a memory and enables learning time dependencies in the input signal. Furthermore, much attention has recently been given to dynamic learning as the latter allows for network adaptation to the changing respiratory patterns of each patient. Indeed, in online training algorithms, the synaptic weights are updated with each new training example. This can be regarded as a way to compensate for irregular movements unseen in the training set and seems promising for medical applications where data is limited. Theoretical advances in online algorithms for training RNNs have recently been made as an attempt to find alternatives to classical approaches (e.g., truncated back-propagation through time), which are subject to shortcomings such as high computational complexity and biased gradient estimation. This thesis aims to examine to which extent the recent algorithmic advances in dynamic training of RNNs enable improving latency compensation for better motion management in lung cancer radiotherapy in the context of limited availability of medical data.

In the first part, internal points close to the lung tumor are tracked in chest computed tomography (CT) scan images using the pyramidal iterative Lucas-Kanade optical flow algorithm. These points may represent fiducial markers, which are small metallic objects implanted in the chest before the radiotherapy treatment to help estimate the tumor position during irradiation delivery. We simultaneously predict their positions using an RNN trained with real-time recurrent learning (RTRL). RTRL recursively computes the influence matrix, which contains the derivatives of the network internal states with respect to the input past signal. The respiratory motion is regular and comprises drift in the spine axis. The sampling rate is approximately equal to 2.5Hz, and the motion amplitude of the tracked points ranges from 12.0mm to 22.7mm. A linear correspondence model enables recovering and predicting three-dimensional (3D) images of a region of interest (ROI) containing the tumor from the positions of the tracked points. The root-mean-square error (RMSE) corresponding to prediction 400ms in advance was equal to 0.44mm, and the average prediction time per time step was approximately 119ms (Dell Intel Core i9-9900K 3.60GHz CPU NVidia GeForce RTX 2080 SUPER GPU). The tumor position in the predicted images appeared visually correct, which is confirmed by the high cross-correlation between the ground-truth and predicted images, equal to 0.955.

In the second part, we study the prediction of the positions of external markers on the chest and abdomen recorded by an infra-red (IR) camera (Polaris system). These IR-reflective objects are used to infer the tumor location using a geometric correspondence model during the treatment. We use nine records of the 3D position of three markers from three healthy individuals breathing during intervals from 73s to 222s. The sampling frequency is 10Hz, the amplitudes of the recorded trajectories vary from 6mm to 40mm in the superior-inferior direction, and almost half of the sequences in the dataset comprise irregular respiratory motion. We simultaneously forecast the location of each marker with a horizon value h (the time interval in advance for which the prediction is made) between 0.1s and 2.0s, using an RNN trained with unbiased online recurrent optimization (UORO). The latter approximates the influence matrix as the product of two random vectors to decrease computational complexity. We compare its performance with RTRL, least mean squares (LMS), and linear regression. Training and cross-validation are performed during the first minute of each sequence. On average, UORO achieved the lowest RMSE and maximum error, equal respectively to 1.3mm and 8.8mm, with a prediction time per time step lower than 2.8ms (Dell Intel Core i9-9900K 3.60Ghz CPU). Linear regression was associated with the lowest RMSE for h = 0.1s and h = 0.2s, followed by LMS for 0.3s ≤ h ≤ 0.6s, and UORO for h ≥ 0.6s.

In the last part, we consider the problem of future frame prediction in chest dynamic magnetic resonance (MR) image sequences. That work is motivated both by the recent development of MRI-guided linear accelerator (LINAC) systems and the interest of the computer vision community in video prediction algorithms. The sampling rate is equal to 3.18Hz, and the respiratory motion is quite regular. Real-time image registration is performed using the pyramidal Lucas-Kanade optical flow algorithm. Principal component analysis (PCA) is used to decompose the time-varying displacement vector field (DVF) into principal (static) deformation fields and low-dimensional time-dependent weights. We compare various algorithms to forecast the latter: RNNs trained with RTRL, UORO, and sparse 1-step approximation (SnAp-1), as well as LMS and linear regression. In SnAp-1, the RNN influence matrix is approximated by a diagonal matrix to reduce computational complexity. Predicting the DVF projection onto the linear PCA feature subspace enables estimating the deformation field in the future and, in turn, the next frames by warping the initial image at $t = t1$ using Nadaraya-Watson regression. Linear regression and LMS respectively led to the most accurate predictions for $h = 0.32s$ and $0.63\text{s} \leq h \leq 1.28\text{s}$. The highest performance for high horizon values, from 1.57s to 2.20s, was achieved by UORO. The corresponding structural similarity index measure (SSIM) between the ground-truth and predicted frames ranged from 0.8978 to 0.8967, and the predicted images were visually close to the original images. The deformation of the organs and the breathing motion also seemed visually correct.

This thesis presents the first applications of RTRL, UORO and SnAp-1 to the prediction of multivariate respiratory signals for safe radiotherapy. It also introduces the first algorithm involving adaptive training of RNNs for dynamic chest MR image prediction. Online training approaches for RNNs performed well for intermediate to high horizon values relative to the sampling frequency. They were robust to irregular respiratory patterns, as suggested by the 10.6% increase in the normalized RMSE between regular and irregular breathing records in the second study, which was the lowest among the methods compared. They were also more stable than LMS, for which hyper-parameter selection with grid search tended to result in high learning rates. Compared with LMS, RNNs resulted in jitter values approximately two times lower. In other words, the predicted signals were not very oscillatory, which makes robotic control of the radiation beam easier. Our work is a step forward in external beam therapy, as it helps accurately estimate the future position of markers and the tumor to compensate for the inherent delay of treatment systems. This will result in higher conformity between the planned and delivered dose, and in turn in a decrease in the occurrence rate and severity of the side effects associated with lung radiotherapy.