Adopting Confident Learning to Eliminate Uncertainty in Chest X-ray Images for Lung Nodules Prediction

Conference Paper · November 2020

DOI: 10.1016/j.ijrobp.2020.07.807

6 authors, including:

Bulat Maksudov
Innopolis University

Ilyas Sirazitdinov
Philips

Ramil Kuleev
Innopolis University

heterogeneous settings using a water and anthropomorphic head-phantom, respectively. The plans were irradiated at our facility and dosimetrically and biologically verified. Furthermore, *in-silico* patient treatment plans were optimized with P, HI, CI as well as $CICR_{CI\text{-}P}$ and $CICR_{CI\text{-}HI}$. Dose-averaged linear energy transfer (LET_d) and RBE with in the target were analyzed and for one case dose to normal tissue surrounding the tumor was investigated.

Results: Measured physical dose differences were ~3 %, while RBE prediction was within 1 %. A biological robustness study for one patient with glioma yielded a similar biophysical stability to P, with $CICR_{CI\text{-}P}$ dose in tissue surrounding the target comparable to that of HI. Median LET_d values in the targets of up to ~30 keVμm^{-1} and ~50 keVμm^{-1} were seen with $CICR_{CI\text{-}P}$ and $CICR_{CI\text{-}HI}$ patient plans. A smaller LET_d, RBE, D_{phys} variability in the target was observed for all CICR plans, compared to SFUD CI treatments. Multi-field $CICR_{CI\text{-}P}$ plans achieved higher homogeneity in RBE, LET_d and D_{phys} distributions.

Conclusion: In this work, we showed that by combining ions in single and multiple fields, more biologically robust and more conformal treatment plans can be delivered. We also performed the first biological and dosimetric verification of multi-ion treatments in a homogeneous and heterogeneous setting.

Author Disclosure: **B. Kopp**: None. **S. Mein**: None. **I. Dokic**: None. **S. Harrabi**: None. **T.T. Böhlen**: None. **T. Haberer**: None. **J. Debus**: Research Grant; CRI (The Clinical Research Institute), View Ray Incl., Accuray International, Accuray Incorporated, RaySearch Laboratories AB, Vision RT limited, Merck Serono GmbH, Astellas Pharma GmbH, Astra Zeneca GmbH, Siemens Healthcare GmbH, Solution Akademie GmbH, Eromed PLC Surrey Research Park, Quintiles GmbH, Pharmaceutical Research. **A. Abdollahi**: None. **A. Mairani**: None.

2762

Adopting Confident Learning to Eliminate Uncertainty in Chest X-ray Images for Lung Nodules Prediction

M. Kholiavchenko,[1] B. Maksudov,[1] I. Sirazitdinov,[1,2] T. Mustafaev,[1,3] R. Kuleev,[1] and B. Ibragimov[1,4]; [1]*Innopolis University, Innopolis, Russian Federation,* [2]*Philips Research, Moscow, Russian Federation,* [3]*University clinic of Kazan Federal University, Kazan, Russian Federation,* [4]*University of Copenhagen, Copenhagen, Denmark*

Purpose/Objective(s): The potential of deep learning to advance lung nodule detection from chest X-rays is significantly compromised by the lack of large annotated databases and noisy labels in the existing databases. The aim of this study is to investigate the applicability of the novel Confident Learning approach for chest X-ray database cleaning and nodule detection improving.

Materials/Methods: We took a subset of the NIH Chest X-ray Dataset of 14 Common Thorax Disease Categories that contains only chest X-ray images with the presence of nodules and the same amount of chest X-ray images of healthy lungs. Next, we split the obtained dataset into train and test sets. In turn, the train set was split into 4-folds to train models using a cross-validation procedure. After that, we trained an Xception (Convolutional Neural Network) model for each fold to classify chest X-ray images with nodules. We calculated probabilities for the whole train set using a cross-validation approach and evaluated the performance of trained models on the test set. To obtain noisy labels, we have to apply a family of theory and algorithms called Confident Learning with provable guarantees of exact noise estimation and label error finding. The algorithm takes noisy labels and predicted probabilities as input and returns found label errors ordered by the likelihood of being an error. We took 5% of the noisiest samples from the list provided by the algorithm and eliminated them from our train set. Then, we repeated the training pipeline but using the clean version of our train set and evaluated it on the test set.

Results: Originally, our classification pipeline gives an accuracy of 71.49%, while after applying the Confident Learning to prune noisy samples, we improved the accuracy to 72.4%. We also brought in a professional radiologist to interpret found label errors by the Confident Learning algorithm. We provided our radiologist with 100 clean chest X-ray images and asked him to classify them. In most cases, radiologist results and known dataset labels for clean X-ray images matched (72%: 36 TP, 36 TN, 18 FP, 10 FN). In this case, false-positive and false-negative predictions by the radiologist can be explained by the fact that the original dataset contains many instances where several pathologies are present in the radiogram at the same time and the radiologist most likely referred these instances to another pathology (not nodular formations). In the next experiment, we gave radiologist 100 noisy chest X-ray images found by the Confident Learning algorithm. It turned out that results obtained from the radiologist and known dataset labels for these noisy X-ray images were very different (only 39% matched: 19 TP, 20 TN, 34 FP, 27 FN).

Conclusion: Our experiments showed that cleaning the datasets can improve the performance of deep learning algorithms on the example of detection of X-rays with lung nodules. Experiments with a radiologist showed that noisy samples are most likely incorrectly labeled or contain atypical cases of nodular formations.

Author Disclosure: **M. Kholiavchenko**: Research Grant; Russian Science Foundation. **B. Maksudov**: None. **I. Sirazitdinov**: Research Grant; Russian Science Foundation. **T. Mustafaev**: Research Grant; Russian Science Foundation. **R. Kuleev**: Research Grant; Russian Science Foundation. **B. Ibragimov**: Research Grant; Russian Science Foundation.

2763

Assessment of Intra-fraction Motion during Frameless Stereotactic Radiosurgery

W. Li,[1,2] G. Bootsma,[3] Y.B. Cho,[2,3] N.J. Laperriere,[2,3] D.B. Shultz,[2,3] and C. Coolens[2,3]; [1]*Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada,* [2]*Department of Radiation Oncology, University of Toronto, Toronto, ON, Canada,* [3]*Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada*

Purpose/Objective(s): Frameless stereotactic radiosurgery (SRS) is facilitated by cone-beam CT (CBCT) imaging, online re-planning, and intra-fraction motion monitoring (IFMM). The purpose of this study was to evaluate the intra-fraction motion detected using frameless immobilization for SRS through the CBCT and IFMM systems.

Materials/Methods: Patients treated with frameless SRS over a 2-month study period were included for analysis. For SRS, patients are immobilized in a personalized thermoplastic mask and headrest followed by attachment of a reflective marker on their nose. The patient's treatment position is verified using a localization CBCT against the reference CBCT acquired at simulation, while the IFMM system continuously monitors their nose marker. Repeat localization CBCTs were acquired during treatment if IFMM thresholds (1.5mm) were exceeded or for voluntary patient breaks. Where feasible, a post-treatment CBCT was acquired after treatment delivery. Using bony anatomy for image registration, the difference between localization and post CBCTs, and localization and repeat localization CBCTs (in situations where IFMM thresholds were exceeded) were used to quantify intra-fraction motion. The IFMM data associated with the time points of the CBCTs were also extracted to assess intra-fraction motion.

Results: Thirty plans were reviewed from 26 patients (19 single fraction, 11 multi-fraction, 4 patients treated concurrently with single and multi-fractionation). Over 52 treatment fractions, 113 localization, repeat localization and post-treatment CBCTs were acquired. 28 sets of localization-post CBCTs and 25 sets of localization-repeat localization CBCTs were included for analysis. The average $\pm$ standard deviation intra-fraction motion quantified on CBCTs was -0.11$\pm$0.37mm, -0.13$\pm$0.20mm, 0.18$\pm$0.74mm in the left/right (L/R), anterior/posterior (A/P) and superior/inferior (S/I) axes. The average $\pm$ standard deviation rotational displacement measured on CBCTs was 0.07$\pm$0.56°, -0.14$\pm$0.64°, 0.02$\pm$0.54° in pitch, yaw and roll. The average $\pm$ standard deviation intra-fraction motion quantified on IFMM was -0.09$\pm$0.68mm, -0.11$\pm$0.5mm, 0.25$\pm$0.76mm in the L/R, A/P and S/I axes. There was a positive correlation between the measured vector CBCT and IFMM displacements ($r = 0.49$).

Conclusion: Preliminary analysis indicate the largest intra-fraction motion, as measured by both CBCT and IFMM systems for frameless GK-SRS, is