

Automated Analysis of CBCT Frequency Using a Polygon-Based Training Model and Random Forest Classification

Inghao Zhou, Arun Gopal, Baoshe Zhang, Huijun Xu, Shifeng Chen, Byong Yong Yi, and Giovanni Lasio,
Department of Radiation Oncology, University of Maryland School of Medicine, Baltimore, MD

Purpose

Purpose: Cone-beam CT (CBCT) is routinely used for image-guided radiation therapy (IGRT); however, in the optimal frequency of CBCT varies across patients and treatment sites, affecting imaging dose, workflow efficiency, and clinical consistency. This study aims to develop an automated machine-learning-based framework to evaluate CBCT acquisition frequency and to assess whether CBCT from the first three treatment fractions can achieve comparable clinical objectives to daily CBCT.

Introduction

Cone-beam CT (CBCT) imaging frequency plays an important role in balancing setup accuracy, imaging dose, and clinical efficiency in image-guided radiation therapy. The abstract is to resolve if it is possible to use the first-three fractions of CBCT to achieve comparable clinical objectives to daily CBCT, particularly in terms of minimizing setup error and enhancing target coverage, which is an extension of a previous paper [1]. There are two novelties in the proposed method. First, A polygon-based geometric representation was employed to encode spatial features derived from CBCT-related datasets. Second, A Random Forest classifier was selected due to its ensemble-based structure, in which multiple decision trees independently generate predictions that are combined through majority voting, improving robustness against noise and overfitting.

Methods and Materials

This Institutional Review Board (IRB)-approved study analyzes 79 lung cancer patients. The in-house IGRT protocol involves daily CBCT for the first three fractions followed by weekly CBCT for soft tissue alignment, along with daily kV-orthogonal for bony anatomy alignment. The eligibility of patient of using this IGRT protocol is determined by a criterion based on numbers of CBCT position matches (equal to 3 fractions match in the first three CBCTs, or great than or equal to 3 fractions match in the first four CBCTs). The fraction matching threshold values (λ) 40–70% of protocol-eligible group (eGroup) were applied, as well as the setup threshold (Δ) values of 3, 4, and 5 mm were applied, respectively.

Sensitivities, specificities and accuracies were computed to quantitate our proposed classification method. A polygon-based geometric representation was employed to encode spatial features derived from CBCT-related datasets (Figure 1 and 2). This approach enables consistent handling of geometric information while preserving clinically relevant spatial relationships. The examples of data applied to the polygon-based training data are shown in Figure 3. A Random Forest classifier was implemented using feature bagging which significantly decreasing the correlation between each decision tree and increasing its predictive average accuracy.

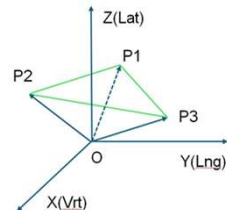


Figure 1 Schematic illustration of the polygon-based training model used to encode geometric features for CBCT frequency classification. P1, P2 and P3 are the CBCT offsets of the first three fractions of treatment. X, y, z are vertical, longitudinal, and lateral directions.

Feature type	Definition	# Elements
three points in 3D space	vertical, longitudinal, lateral of the shift of each CBCT fraction	9
three points in 3D	magnitudes of three points	3
Edge	6 edges of a tetrahedron	6
Angle	12 angles of a tetrahedron	12
Perimeter of triangle	4 faces of a tetrahedron	4
Area of triangle	4 faces of a tetrahedron	4
Total		38

Figure 2. Features are extracted the CBCT offsets of the first three fractions of treatment.

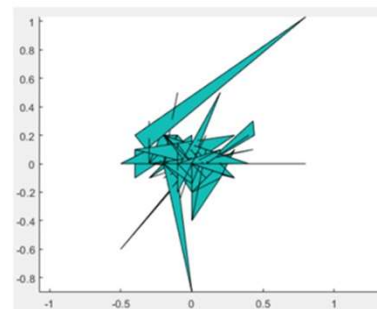


Figure 3 Examples of data applied to the polygon-based training data. Each triangle shows a triangle of P1P2P3 in Figure 1.

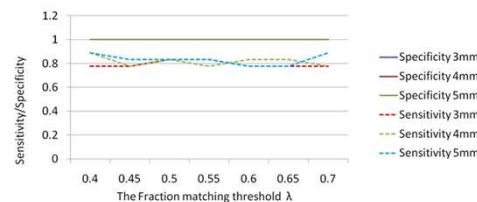


Figure 4 The sensitivity, specificity curves with the fraction matching threshold of ground truth of 40%-70% with step size of 5%, as well as the setup threshold (ST) of 3, 4 and 5mm.

Results

The sensitivity and specificity curves with λ from 40% to 70% with step size of 5%, as well as Δ values of 3, 4, and 5 mm are shown in Figure 4. When λ equals 50%, sensitivities for identifying a protocol-ineligible group (IGroup) patient were 0.83, 0.83, and 0.83, specificities for identifying an eGroup patient were 1.0, 1.0, and 1.0, respectively.

Conclusions

The combination of polygon-based feature representation and Random Forest classification provides an accurate and interpretable framework for CBCT frequency analysis which shows it's possible to use the first-three fractions of CBCT to achieve comparable clinical objectives.

References

[1] Zhou, J., Gopal, A., Zhang, B., Xu, H., Chen, S., Yi, B. and Lasio, G., 2025. A novel optimization of cone beam CT frequency for lung radiation therapy based on an image-guided radiation therapy protocol and patient classification method. Translational Lung Cancer Research, 14(1), p.81.