2026 JOINT AAPM | COMP MEETING
JULY 19–22 | VANCOUVER, BC

Predicting the Necessity of Confirmation CBCT in Lung IGRT Using Machine Learning Classification

Sophia Zhang[1]; Chengyu Shi, PhD[2]
[1] Irvine High School, [2]City of Hope Department of Radiation Oncology

2026 JOINT AAPM | COMP MEETING
JULY 19–22 | VANCOUVER, BC

Abstract

Standard kV/kV pairs offer fast, low-dose alignment in lung cancer radiation therapy. This study utilizes machine learning to analyze 1,063 patients to determine which clinical factors influence the critical need for confirmation CBCT. The MATLAB app was used to train 25 different machine learning models to predict post-alignment anatomical shifts based on clinical variables. An ANOVA ranking algorithm calculated importance scores to isolate drivers of these setup variations. The models resulted in 84.1% accuracy with Efficient Linear SVM peaking at 85.77%. Post-alignment shifts correlate most strongly with higher T and N tumor stages, advanced patient age, male patients, larger PTV volumes, and right lung treatment sites. Post-alignment anatomical shifts are highly predictable based on the patient's clinical profile. These data-driven findings directly improve personalized IGRT protocols that allow clinical teams to selectively utilize CBCTs to balance tumor target precision with minimized treatment dose and times.

Introduction

- The precise alignment of patients in Image-Guided Radiation Therapy (IGRT) is crucial for setup variations and protecting surrounding healthy tissue
- While standard kV/kV pairs offer fast, low-dose alignment, it remains unclear if 3D confirmation CBCT is always necessary given its extra treatment time and imaging dose
- **The purpose of this project is to utilize machine learning classification applications in MATLAB to analyze 1,063 lung patient fractions and determine which clinical predictors (age, sex, tumor stage, PTV volume) dictate the need for a confirmation CBCT**

Methods and Materials

- Conducted a study analyzing a dataset of 1,063 lung cancer patient treatment fractions
- Utilized MATLAB Classification Learner app to train and evaluate 25 distinct machine learning models (such as SVMs and Trees) in order to predict post-alignment shifts
- Applied an ANOVA ranking algorithm to calculate feature importance scores across seven clinical variables:
  - Tumor Metrics: T stage, N stage, PTV volume, and treatment site
  - Patient Metrics: Age, sex, and specific kV/kV-pair adjustments used

Results

- Across 25 MATLAB classification models, the system achieved 84.1% accuracy in predicting post-alignment shifts, proving that post-alignment shifts are predictable rather than random
- ANOVA feature demonstrated that the need for confirmation CBCT is dictated by specific clinical factors:
  - PTV Volume and Tumor Stage: Larger volume and higher T/N stages show strongest correlatio with significant shifts
  - Patient Demographics: Older patients and male patients were statistically easier to shift post-kV/kV alignment
  - Tumor Location: Tumors in the right lung showed higher frequency of shifting than tumors in the left lung

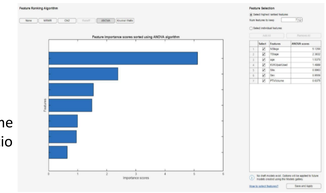

Figure 3. Predictors' ANOVA ranking

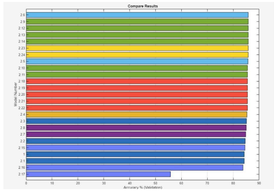

Figure 1. Result comparison for different models

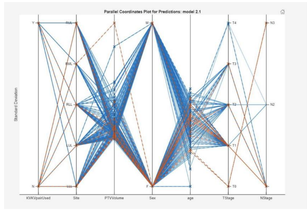

Figure 2. Parallel plots for predictors

Table 1. Matlab applications classification models comparison.

Model Number	Model Type	Status	Accuracy % (Validation)	Total Cost (Validation)
1	Tree	Trained	84.30962	150
2.1	Tree	Trained	84.30962	150
2.2	Tree	Trained	84.82343	147
2.3	Tree	Trained	85.14644	142
2.4	Binary GLM Logistic Regression	Trained	85.35565	140
2.5	Efficient Logistic Regression	Trained	85.66946	137
2.6	Efficient Linear SVM	Trained	85.77406	136
2.7	Naive Bayes	Trained	84.83264	145
2.8	Naive Bayes	Trained	85.04184	143
2.9	SVM	Trained	85.77406	136
2.1	SVM	Trained	85.56485	138
2.11	SVM	Trained	85.56485	138
2.12	SVM	Trained	85.77406	136
2.13	SVM	Trained	85.77406	136
2.14	SVM	Trained	85.77406	136
2.15	Ensemble	Trained	84.51883	148
2.16	Ensemble	Trained	83.78661	155
2.17	Ensemble	Trained	55.75314	423
2.18	Neural Network	Trained	85.56485	138
2.19	Neural Network	Trained	85.56485	138
2.2	Neural Network	Trained	85.56485	138
2.21	Neural Network	Trained	85.56485	138
2.22	Neural Network	Trained	85.56485	138
2.23	Kernel	Trained	85.77406	136
2.24	Kernel	Trained	85.77406	136

Figure 4. Matlab applications classification models comparison

Discussion

- **Anatomical Drivers of Setup Errors**: ANOVA ranking identified tumor staging (N-stage: 5.12, T-stage: 2.38) as the primary drivers of post-alignment setup variations. Advanced stages typically present with larger tumor volumes, increased nodal development, and complex structural changes, which interfere with accuracy of standard 2D alignments.
- **Demographic and Physiological Variances**: Advanced patient age and male sex correlate with an increased likelihood of anatomical shifts. This is associated with age-related breathing irregularities and difficulties maintaining a stable position on the treatment table.
- **Tumor Location Bias**: Right-lung tumors exhibited higher probability for post-alignment shifting compared to left-lung tumors. This is likely due to the structurally larger volume and mobility of the right lung.
- **Clinical Protocol Optimization**: Since these shifts are highly predictable rather than random, clinical teams can utilize these specific factors as a baseline screening checklist. High-risk patients can selectively be flagged for mandatory 3D CBCT scanning, while low-risk patients can safely skip it.

Conclusions

Machine learning analysis of 1,063 lung patient fractions achieved an average accuracy of 84.1% across 25 MATLAB models, proving that CBCT confirmation remains necessary. Post-kV/kV alignment shifts correlate with larger PTV volumes, older age, male patients, higher T/N tumor stages, and right lung treatment sites.

These findings provide data-driven insights to optimize personalized IGRT protocols. By predicting which patients are most prone to shifting, clinical teams can precisely use CBCT confirmation to balance maximum tumor precision with minimal additional imaging dose and treatment time.

Contacts

Chengyu Shi, PHD, DABR
City of Hope, Radiation Oncology
cshi@coh.org

Sophia Zhang
Irvine High School
sophiazijunz@gmail.com