JULY 12–16 2020 VIRTUAL JOINT AAPM | COMP MEETING
EASTERN TIME [GMT-4]

# Variations in radiomics features of a multi-texture phantom introduced by deep learning iterative reconstruction algorithms

N. BAUGHAN[1], J. P. CRUZ-BASTIDA[1], H. AL-HALLAQ[1], and I. REISER[1]

[1]University of Chicago, Chicago, IL

AT THE FOREFRONT UChicago Medicine

## INTRODUCTION

- Recently developed CT deep learning iterative reconstruction (DLIR) algorithms aim to improve image quality over currently established algorithms such as filtered back projection (FBP) and adaptive statistical iterative reconstruction (ASIR).[1]
- However, the impact of DLIR algorithms in a variety of media on quantitative image metrics, such as radiomic features, may result in feature values which deviate from those found with established algorithms, and as a result, could correlate with the performance of automated lesion classifiers.

## AIM

**Perform** radiomic analysis of CT images of a six-texture phantom reconstructed with five reconstruction algorithms, for two reconstruction kernels.

**Evaluate** the variations in radiomics features due to the five different noise reduction algorithms.

**Assess** the magnitude of variations introduced by each noise reduction algorithm in comparison to the magnitude of variation introduced by doubling the scan pixel size.

## METHOD

| CT scans of the updated Credence Cartridge Radiomics phantom[2] | |
|---|---|
| Reconstruction kernel | Lung and standard soft-tissue |
| Pixel sizes | 0.586 mm and 0.293 mm |
| Reconstruction algorithms | FBP, ASIR (50%), DLIR high, DLIR medium, and DLIR low |
| Tube current | 200 mAs |
| Tube voltage | 120 kV |
| Scan dose | 13.60 mGy |
| Slice thickness | 2.5 mm |

- Radiomics analysis performed with open-source software PyRadiomics[3]
  - A cylindrical region of interest was selected in the center of each of the six texture regions, 7.5 cm in diameter and 1.5 cm thick
  - PyRadiomics was used to extract first-order and GLCM texture features
  - Calculated differences between the two pixel size scans for each of the reconstruction algorithms for all textures

## RESULTS

- Figure 1 shows representative CT slices of the six textures in the phantom, developed by Ger *et al.*[2], from which texture features were extracted.
- Figure 2 shows variations in radiomics features across five noise-reduction settings for two reconstruction kernels. Vertical bars indicate the differences in feature values caused by doubling the pixel size. Symbols, noted in the legend, specify the region of the phantom from which features were extracted.
- Feature values from FBP and ASIR or the three levels of DLIR tended to cluster together, and trends were similar across all texture regions.
- For the standard soft-tissue kernel, variations due to the reconstruction algorithm tended to be larger than the variation of changing the pixel size by a factor of two, for both first-order and gray level co-occurrence (GLCM) features.
- For the lung kernel, feature value variations due to changing the pixel size by a factor of two tended to be larger than variations due to the reconstruction algorithm.

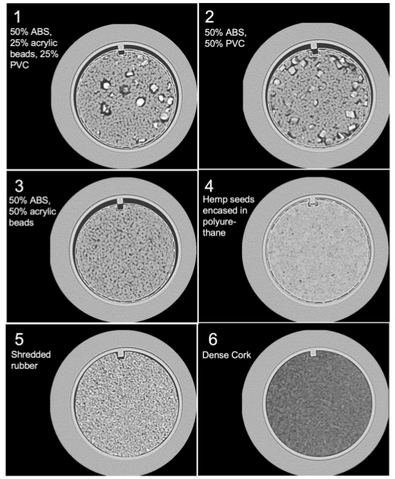


*Figure 1. Representative slices of the six texture regions in the phantom, shown with the lung filter, reconstructed with FBP. Material percentages by weight.*

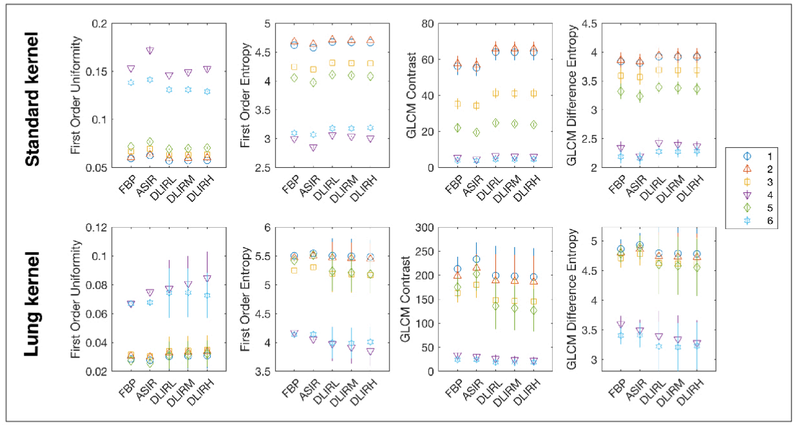


*Figure 2. First-order and GLCM feature values on an absolute scale, but with error bars which represent the difference in feature values when doubling the pixel size from 0.293 mm to 0.586 mm for filtered back projection (FBP), adaptive statistical iterative reconstruction (ASIR), and deep learning iterative reconstruction (DLIR) low (L), medium (M), and high (H).*

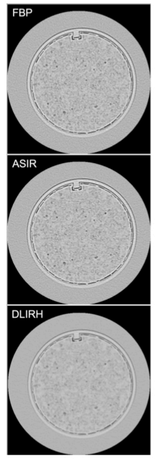


*Figure 3. Texture region 4 shown for three reconstruction algorithms with the lung filter.*

## CONCLUSIONS

- For the selected features, there was little variation between established algorithms FBP and ASIR.
- However, reconstruction algorithms have the potential to cause variations of radiomics features larger than those caused by doubling the pixel size in a variety of textures.
- Variations found between established algorithms and DLIR algorithms indicate a potential need to consider standardization when applying additional automated classification.
- Future work on standardization of quantitative image metrics between reconstruction algorithms, similar to the work done by Mackin *et al.*, indicate a potential solution for use of these metrics in computer aided diagnosis programs.[4]

## REFERENCES

[1] J. Hsieh, E. Liu, B. Nett, J. Tang, J. Thibault, and S. Sahney, A new era of image reconstruction : TrueFidelity; Technical white paper on deep learning image reconstruction, (2019).

[2] R.B. Ger, S. Zhou, P.C.M. Chi, *et al.*, Comprehensive Investigation on Controlling for CT Imaging Variabilities in Radiomics Studies, Sci. Rep. **8**(1), 1–14 (2018).

[3] J.J.M. Griethuysen, A. Fedorov, C. Parmar, C., *et al.*, Computational Radiomics System to Decode the Radiographic Phenotype, Cancer Research, 77(21), e104–e107 (2017).

[4] Mackin, D., Ger, R., Dodge, C. *et al.* Effect of tube current on computed tomography radiomic features. *Sci Rep* **8,** 2354 (2018).

## ACKNOWLEDGEMENTS

The authors would like to thank the Court Lab for the use of the updated Credence Cartridge Radiomics phantom, and K. Blunt and V. Mamyan for scanning the phantom.

This work was supported in part by the NIH T-32 Training Grant T32 EB002103. H.A.A. received royalties and licensing fees for computer-aided diagnosis technology through the University of Chicago.

## CONTACT INFORMATION

Natalie Baughan: nbaughan@uchicago.edu

Ingrid Reiser: ireiser@uchicago.edu