A Quantification Method for Comparing Low-Contrast Edge Definition on Clinical Head CT

Nicholas K Mathew, MS[1] , Brian E. Nett PhD[1] , Gregory D Avey, MD[2] , Tabassum A Kennedy, MD[2] , Aaron S Field, MD[2] , Dr. Timothy P. Szczykutowicz, PhD PhD[3] and Aria M. Salyapongse [3] ,

(1) GE Healthcare Technologies, Waukesha, WI, (2) University of Wisconsin School of Medicine and Public Health, Madison, WI, (3) Departments of Radiology and Medical Physics, University Wisconsin-Madison, Madison, WI

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

ABSTRACT

Purpose: To propose a clinically relevant, quantitative comparison of low-contrast edge definition on routine CT Head examinations.

Methods: Routine head CT exams were collected from an investigational photon-counting CT and an energy-integrating CT. The exams were processed with baseline reconstruction settings typically used in clinical practice. Three radiologists were tasked to create six ROIs over specific boundaries between anatomies in the Alberta Stroke Program Early CT Score over each scan volume. A Python-based tool was developed to facilitate ROI drawing by using a paintbrush-like ROI drawer. The measurement, termed "Blur Area", was defined as the number of pixels within the ROI with a value within the 25th and 75th percentile range. This metric was calculated on the baseline reconstruction and additional reconstructions to compare the effect of different parameters on the low-contrast edges defined by the ROIs.

Results: ROIs were collected and the blur area was computed for the baseline image and three different reconstruction algorithms. We saw an overall decrease in the blur area across all algorithms in comparison to the baseline. Blur area percentage change showed high variance at small ROI sizes, stabilizing at larger ROI sizes indicating that ROI maximization and sampling should be maximized for better precision of the measurement.

Conclusions: This work demonstrates a method for quantifying the impact of algorithms on the image quality of clinical images to serve their clinical task. In the framework presented, a radiologist manually asserted the areas analyzed in their clinical assessment to establish the most clinically relevant areas of the image. This method is more relevant than phantom measurements as it operates on clinical images which are inherently subject to the relevant conditions that may affect low-contrast edge measurements such as noise and contrast.

CONTACT

Nick Mathew
GE HealthCare (Computed Tomography)
3000 N Grandview Blvd, Waukesha WI
nick.mathew@gehealthcare.com

INTRODUCTION

Assessment of low-contrast detectability (LCD) in computed tomography has traditionally relied on standardized phantom analysis, which provides reproducible measurements of noise, contrast, contrast-to-noise ratio (CNR), and LCD. Though these methods are essential for system benchmarking and quality assurance, they do not fully reflect the complexity of clinical imaging, where image quality depends on interacting factors such as radiation dose, reconstruction algorithm, noise magnitude, noise texture, spatial resolution, object size and contrast, and anatomical background structure - all of which affect the visibility and sharpness of subtle low-contrast boundaries. Prior work has highlighted the limitations of phantom-only evaluations and the need for clinically relevant, task-based image quality assessment [1-2]. Similar software-assisted image analysis has been applied in [3] in which cortical and trabecular bone delineation in wrist CT scans were assessed via analysis of the interquartile range, corresponding to the blur between the tissues. Motivated by the need for clinical-based image quality assessment, we developed a quantitative method to compare low-contrast edge definition in clinical head CT images, with the goal of enabling more meaningful assessment of image quality in the context of diagnostic interpretation.

METHODS AND MATERIALS

Routine head CT exams were collected from a photon-counting CT (Photonova Spectra, GE HealthCare) and an energy-integrating CT (Revolution Apex, GE HealthCare). The exams were processed with typical reconstruction settings for routine head CT at 22cm DFOV with 2.5mm slice thickness. Three radiologists were tasked to create ROIs over six specific anatomy boundaries (Table 1) in the Alberta Stroke Program Early CT Score [4] over each scan volume. A Python-based tool was developed to facilitate ROI drawing by using a paintbrush-like ROI drawer. The measurement, termed "Blur Area", was defined as the number of pixels within the ROI with a magnitude within the 25th and 75th percentiles. This metric was calculated on the baseline reconstruction and three progressively higher resolution reconstructions to compare the sharpening effect on clinically relevant low-contrast boundaries.

Table 1: The anatomy boundaries ROI targets used.

Boundary Target Type (Anatomy 1 – Anatomy 2)	Anatomy 1	Anatomy 2
Gray Matter – White Matter	Putamen	Internal Capsule
	Posterior MCA Cortex (1 gyrus)	Bordering White Matter
	Anterior MCA Cortex (1 gyrus)	Bordering White Matter
Gray Matter – CSF	Caudate Head	Lateral Ventricle Anterior Horn
	Cortex Lateral to insular ribbon	Sylvian Fissure
CSF – White Matter	Posterior Horn Lateral Ventricle	Bordering White Matter

RESULTS

ROIs were collected and the blur area was computed for the baseline image and three different reconstruction algorithms. We saw an overall decrease in the blur area across all algorithms in comparison to the baseline reconstruction. Blur area percentage change showed high variance at small ROI sizes, stabilizing at larger ROI sizes indicating that ROI maximization and sampling should be maximized for better precision of the measurement.

Table 2: Median and Average change in Blur Area across all ROI's, stratified by radiologist and reconstruction type.

	Radiologist 1			Radiologist 2			Radiologist 3		
Reconstruction Type	1	2	3	1	2	3	1	2	3
Median Δ (in %)	-7.14	-15.48	-22.58	-6.30	-13.44	-18.60	-7.74	-14.76	-21.21
Average Δ (in %)	-7.12	-14.29	-21.06	-6.29	-13.36	-18.68	-7.35	-15.10	-21.09

Figure 1: Summary of all blur area changes grouped by reconstruction type.

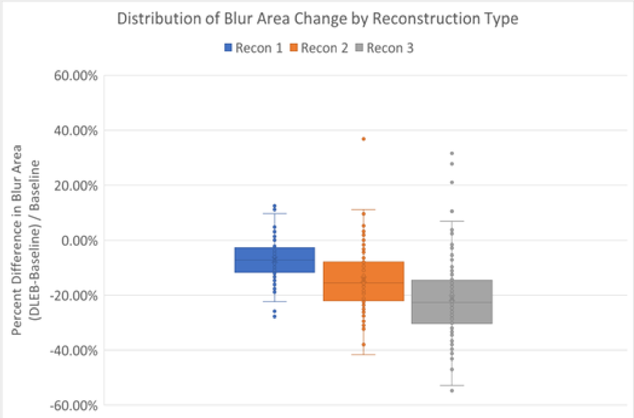


Figure 2: All data points plotted against the size of the drawn ROI.

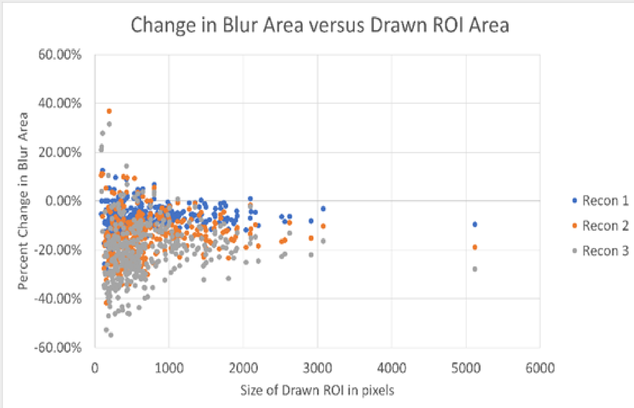
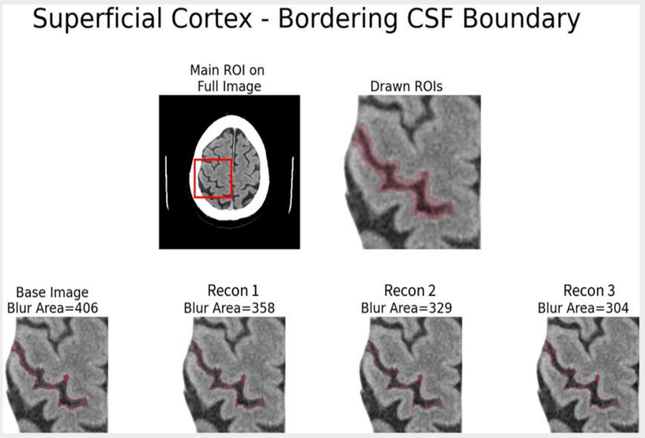


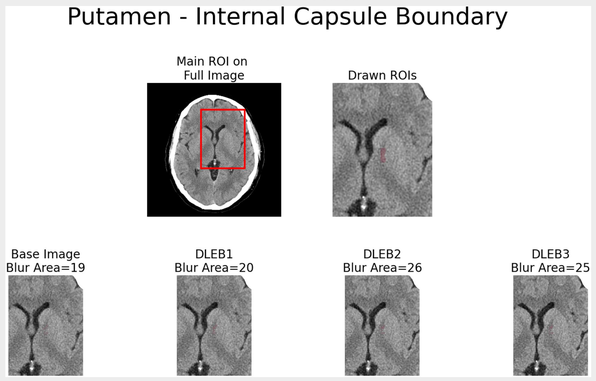
Figure 3: Example of a drawn ROI and the corresponding calculated blur areas for each reconstruction type.



DISCUSSION

The results aligned with the expected progressive shrinking of the Blur Area. While the trend was observed across each radiologist, there was considerable variation in the magnitude of the change. In addition, there was also variation that correlated to the size of the ROI. A similar study in [5] analyzed the effect of ROI size in the ability to assess abnormalities in renal tissues, finding reduced confidence with smaller ROIs. Outliers existed in the dataset, which indicated large changes in the blur area. Further analysis showed outliers to be correlated with very small ROIs, further emphasizing the effect of ROI size.

Figure 4: Outlier corresponding to large percentage increase, driven by small ROI size.



CONCLUSIONS

This work demonstrates a method for quantifying the impact of algorithms on the image quality of clinical images to serve their clinical task. In the framework presented, a radiologist manually asserted the areas analyzed in their clinical assessment to establish the most clinically relevant areas of the image. This approach can be applied to analyze boundary delineation at radiologist defined ROIs. This method is more relevant than phantom measurements as it operates on clinical images which are inherently subject to the relevant conditions that may affect low-contrast edge measurements such as noise and contrast. The main limitation of the study is the comparison of only reconstruction techniques that do not affect the spatial grid. Future work should assess the method in its ability to maintain quantitative accuracy in unregistered formats.

REFERENCES

[1] E. Samei *et al.*, "Performance evaluation of computed tomography systems: Summary of AAPM Task Group 233," *Medical Physics*, vol. 46, no. 11, pp. e735–e756, 2019, doi: 10.1002/mp.13763.
[2] T. B. Smith, J. B. Solomon, and E. Samei, "Estimating detectability index in vivo: development and validation of an automated methodology," *JMI*, vol. 5, no. 3, p. 031403, Dec. 2017, doi: 10.1117/1.JMI.5.3.031403.
[3] J.-P. Grunz *et al.*, "Evaluation of Ultra-High-Resolution Cone-Beam CT Prototype of Twin Robotic Radiography System for Cadaveric Wrist Imaging," *Academic Radiology*, vol. 28, no. 10, pp. e314–e322, Oct. 2021, doi: 10.1016/j.acra.2020.06.018.
[4] N. O. Pop *et al.*, "The Alberta Stroke Program Early CT score (ASPECTS): A predictor of mortality in acute ischemic stroke," *Exp Ther Med*, vol. 22, no. 6, p. 1371, Dec. 2021, doi: 10.3892/etm.2021.10805.
[5] M. S. Davenport, A. M. Neville, J. H. Ellis, R. H. Cohan, H. S. Chaudhry, and R. A. Leder, "Diagnosis of Renal Angiomyolipoma with Hounsfield Unit Thresholds: Effect of Size of Region of Interest and Nephrographic Phase Imaging," *Radiology*, vol. 260, no. 1, pp. 158–165, Jul. 2011, doi: 10.1148/radiol.11102476.

© Poster Template by Genigraphics® 1.800.790.4001 www.genigraphics.com