

# Comparison of Two Alternative Forced Choice (2-AFC) and Receiver Operating Characteristic (ROC) Methods for Low Contrast Detectability Using the Cathpan 503 Phantom

Yanwar Budhiyanto<sup>1</sup>, Eko Hidayanto<sup>1</sup>, Catur Edi Widodo<sup>1</sup>, Choirul Anam<sup>2</sup>

<sup>1</sup>Department of Physics, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia.

Corresponding Author: Yanwar Budhiyanto

DOI: <https://doi.org/10.52403/ijrr.20260346>

## ABSTRACT

Detecting small and low-contrast objects in *CT-Scan* images remains a major challenge in radiological diagnostic practice. The limitations of human visual capabilities and the influence of image quality require objective and reliable quantitative evaluation methods. The effect of object size variation on *CT-Scan* image detection performance was analyzed using the *Two Alternative Forced Choice (2-AFC)* and *Receiver Operating Characteristic (ROC)* psychophysical approaches.

The study was conducted using a *Catphan 503 phantom* with low-contrast objects measuring 6 mm, 8 mm, and 10 mm in diameter. Image acquisition was performed using a Toshiba Aquilion *CT-Scan* machine with parameters of 120 kV, 300 mA, axial mode, and a slice thickness of 2 mm. The scanned images were processed using *IndoQCT* software to generate *2-AFC* test stimuli and converted to *ROC* format. Measurements were performed using *Region of Interest (ROI)* on the object and background areas. Performance evaluation was conducted by five medical physicist observers. Analysis parameters included sensitivity, specificity, and *Area Under Curve (AUC)*.

This study concluded that there was a consistent correlation between the results of the *2-AFC* and *ROC* methods, indicating that the two methods complement each other in evaluating the detection performance of *CT-Scan* images. Object size had a significant effect on the accuracy of visual detection in *CT-Scan* images. The combined approach of the *2-AFC* and *ROC* methods is effective as a quantitative evaluation tool for the detectability of low-contrast objects.

**Keywords:** *CT-Scan*, object size, *2-AFC*, *ROC*, low-contrast detectability

## INTRODUCTION

Computed Tomography (CT) is a primary diagnostic imaging modality capable of producing detailed anatomical representations through X-ray-based tomography reconstruction. Diagnostic accuracy is greatly influenced by image quality, particularly the system's ability to detect small and low-contrast objects. Failure to detect at this stage has the potential to cause clinical errors that directly impact patient safety.

Various *CT-Scan* Quality Control (QC) studies show that even though the system meets regulatory standards, variations in CT number values, uniformity, and spatial

resolution remain major challenges, especially in the detection of small objects. This problem becomes even more complex when human visual limitations and image noise play a role in the interpretation pROcess.

Psychophysical approaches such as Two-Alternative Forced Choice (2-AFC) and Receiver Operating Characteristic (ROC) have been widely used to evaluate medical image detection performance. The 2-AFC method directly measures sensitivity through forced choice between two images, while ROC evaluates the balance between sensitivity and specificity through AUC analysis. However, studies that integrate these two methods with a primary focus on object size variation in CT-Scan images are still limited. Therefore, this study focuses on analyzing the effect of object size on CT-Scan image detection performance by combining the 2-AFC and ROC methods quantitatively and systematically.

## RESEARCH METHOD

### Research Design

This study is a quantitative experimental study using the *Catphan 503 phantom* as the test object.

### Image Acquisition

Images were obtained using a Toshiba Aquilion CT-Scan with the following parameters: 120 kV, 300 mA, slice thickness 2 mm, axial mode, and revolution time 0.75 s. The phantom was positioned exactly at the center of the gantry to minimize geometric artifacts.

### Objects and ROI

Low-contrast objects with diameters of 6 mm, 8 mm, and 10 mm were analyzed using two ROIs: a green object ROI and a red background ROI. The size of the ROI was adjusted to the diameter of each object.



Figure 3.4 ROI Creation process in IndoQCT

Tabel 4.1 ROI Size Adjusted to Object Diameter

No	Diameter Object	ROI Size (mm)
1	6 mm	10.73
2	8 mm	11.99
3	10 mm	13.25

### 2-AFC and ROC Tests

Each object size was tested 100 times using 2-AFC. The data was then converted into 200 ROC data points with five confidence levels. The evaluation was performed by five experienced medical physicists

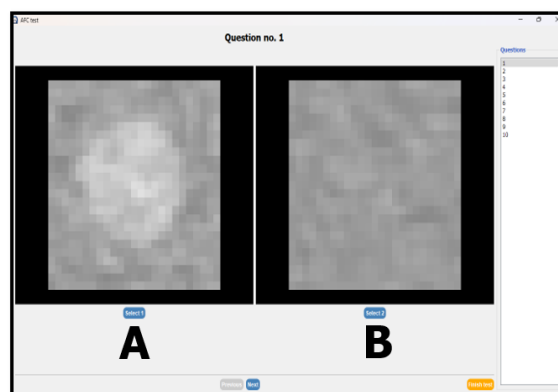


Figure 3.5 2-AFC Test



Figure 3.6 ROC Test

## Data Analysis

Detection performance was analyzed using the percentage of correct answers (2-AFC), sensitivity, specificity, and ROC AUC. The relationship between 2-AFC and ROC was analyzed descriptively and comparatively

## RESULT

The results of the 2-AFC test show an increase in the percentage of correct answers as the object size increases.

**Table 4.2 Results of 2-AFC Test Measurements**

No	Size Object	Average Correct Answer
1	6 mm	76%
2	8 mm	96%
3	10 mm	99%

ROC analysis shows a consistent increase in AUC values

**Table 4.3 Results ROC Test Measurements**

No	Size Object	Average AUC
1	6 mm	0.76
2	8 mm	0.80
3	10 mm	0.81

These results show a linear relationship between object size, sensitivity, and overall diagnostic performance.

## DISCUSSION

The 2-AFC and ROC test results in Tables 4.2 and 4.3 indicate an improvement in object detection performance as object size increases from 6 mm to 10 mm. In the 2-AFC test, the percentage of correct answers jumped from 76% to 99%, while the AUC value in the ROC analysis increased from 0.76 to 0.81. This trend reflects a positive relationship between object size and observer sensitivity, where larger objects make it easier to distinguish signals from noise in medical images such as CT scans. This improvement is consistent with the basic principle of spatial resolution in imaging systems, where object sizes approaching the imaging resolution limit will decrease detection accuracy.

The most significant increase in the percentage of correct answers in 2-AFC

occurred from 6 mm (76%) to 8 mm (96%), followed by a more gradual increase to 10 mm (99%). This indicates a detection threshold of around 8 mm, at which observers achieved near-perfect performance. In the context of 2-AFC, which measures the observer's ability to distinguish objects from a blank background, these results underscore the influence of size on the signal-to-noise ratio (SNR). Small objects such as 6 mm are susceptible to the partial volume effect and quantum noise, thereby reducing the observer's confidence in selecting the correct alternative.

ROC analysis complements the 2-AFC findings by measuring overall diagnostic performance through AUC. The AUC value is 0.76 at 6 mm, while it is 0.80–0.81 at larger sizes. The relatively small increase in AUC from 8 mm to 10 mm (0.80 to 0.81) indicates performance saturation, where the limiting factor shifts from object size to imaging system limitations or observer variability. The resulting ROC curve likely shows an optimal trade-off between sensitivity and specificity at higher thresholds for larger objects

## CONCLUSION

Object size significantly affects CT-Scan image detection performance. The combination of the 2-AFC and ROC methods proved effective in evaluating the detection quality of small and low-contrast objects. This approach is recommended as a quantitative evaluation method in the optimization of clinical CT-Scan protocols.

### Declaration by Authors

**Acknowledgement:** None

**Source of Funding:** None

**Conflict of Interest:** No conflicts of interest declared.

## REFERENCES

1. Abbey, C.K., Eckstein, M.P., & Bochud, F.O. (1999). Estimation Of Human-Observer Templates In Two-Alternative Forced-Choice Experiments. *Medical Imaging*.

2. Al-Naser, Y., & Tafti, D. (2023). *Ct Instrumentation And Physics. Statpearls.*
3. Anam, C., Amilia, R., Naufal, A., & Dougherty, G. (2023). Automatic Measurement Of *Ct* Number Linearity In Three Types Of Cathpan Phantoms. *Journal Of Technology.*
4. Asni, N., & Utami, M. S. N. (2023). *Ct Scan Quality Control (Image Quality Analysis And Evaluation). Metode, 4, 13*
5. Write Relevant, Correct And Complete References Here.
6. Bushberg, J. T., Seibert, J. A., Leidholdt, E. M., Boone, J. M. (2011) . *The Essential Physics Of Medical Imaging.* (N.P.): Wolters Kluwer Health.
7. Champley, K.M., Zellner, M.B., Tringe, J.W., & Martz, H.E. (2024). Methods For Few-View *Ct* Image Reconstruction. *Arxiv, Abs/2410.07552.*
8. Choopani, M., Abedi, I., & Dalvand, F. (2023). Quality Assessment Of *Computed Tomography* Images Using A Channelized Hotelling Observer: Optimization Of Protocols In Clinical Practice. *Advanced Biomedical Research, 12.*
9. Deng, L., Feng, P., Chen, M., He, P., Quang-Sang, V.Q., & Wei, B. (2014). A *Ct* Reconstruction Algorithm Based On Non-Aliasing Contourlet Transform And Compressive Sensing. *Computational And Mathematical Methods In Medicine, 2014.*
10. Dewantari, R., Anam, C., Sutanto, H., Naufal, A., Amilia, R., Izmi, S.I., Putri, H.S., Dewi, P.S., Ilham, I.R., Haryanto, F., & Setiawan, A.M. (2024). *2-AFC* For Detectability Of Low Contrast *Objects* In *Ct* Images Scanned With Two Doses And Reconstructed With Various Iterative Reconstruction (Ir) Levels. *International Journal Of Scientific Research In Science And Technology.*
11. Gao, Y., Jiang, Y., Peng, Y., Yuan, F., Zhang, X., & Wang, J. (2025). Medical Image Segmentation: A Comprehensive Review Of Deep Learning-Based Methods. *Tomography, 11.*
12. Genske, U., & Jahnke, P. (2022). Human Observer Net: A Platform Tool For Human Observer Studies Of Image Data. *Radiology, 303(3), 524-530.*
13. Gideon, S. (2015). Simulation-Based *Ct* Imaging Learning Using Image Reconstruction Principles In Matlab Software.
14. Grose, J., & Pick, G. (2024). The Calibration And Validation Of A Two-Alternative Forced-Choice Test For Evaluating Hearing Loss Of Cochlear Origin. *Autumn Conference 1979.*
15. Handbook Of X-Ray Imaging: Physics And Technology. (2017). United Kingdom: Crc Press.
16. Hendee, W. R., Ritenour, E. R. (2003). *Medical Imaging Physics.* Germany: Wiley.
17. Hsieh, J. (2022). *Computed Tomography: Principles, Design, Artifacts, And Recent Advances, Fourth Edition.*
18. Hsieh, J. (2003). *Computed Tomography: Principles, Design, Artifacts, And Recent Advances.* United States: Spie Press.
19. Ilham, I.R., Anam, C., Sutanto, H., Naufal, A., & Amilia, R. (2024). Impact Of Radiation Dose And Iterative Reconstruction (Ir) Level On Low-Contrast Detectability With 4-Afc Approach. *International Journal Of Scientific Research In Science And Technology.*
20. Inayah, S.N., Anam, C., Sutanto, H., Naufal, A., & Amilia, R. (2025). Comparative Analysis Of Low-Contrast Detectability (Lcd) Using A 4-Afc: Filtered Back Projection (Fbp) And Iterative Reconstruction (Ir) Images. *International Journal Of Scientific Research In Science And Technology.*
21. Jabbar, A., Universitas, L., Medan, H.K., No, J.H., & Medan, C.S. (2020). Utilization Of *CT-Scan* (Computer Tomography) In The Medical Field.
22. Jung, H. (2021). Basic Physical Principles And Clinical Applications Of *Computed Tomography.* *Progress In Medical Physics, 32(1), 1-17.*
23. Kemp, T.M., Kawakita, T.A., Mehta, R., & Pineda, A.R. (2023). Optimizing Data Acquisition In Undersampled Magnetic Resonance Imaging (Mri) Using *Two Alternative Forced Choice (2-AFC)* And Search Tasks. *Medical Imaging.*
24. Lee, H.S., Hout, D.V., & Hautus, M.J. (2007). Comparison Of Performance In The A-Not A, *2-AFC*, And Same-Different Tests For The Flavor Discrimination Of Margarines: The *Effect* Of Cognitive Decision Strategies. *Food Quality And Preference, 18, 920-928.*
25. Li, Y., Jiang, Y., Liu, H., Yu, X., Chen, S., Ma, D., Gao, J., & Wu, Y. (2021). A Phantom Study Comparing Low-Dose *Ct*

- Physical Image Quality From Five Different Ct Scanners. *Quantitative Imaging In Medicine And Surgery*, 12 1, 766-780.
26. Lockhart, H.A., & Emrich, S.M. (2025). Low-Priority Items Are Held In Visual Working Memory: Evidence From Flexible Allocation In A Two-Alternative Forced-Choice (2afc) Paradigm. *Journal Of Vision*, 25.
  27. Machmudah, Muttaqin, A., & Oktavia, A.B. (2025). Testing The Image Quality Of Ct Scans Using Acr (American College Of Radiology) Phantoms In The Radiology Department Of Andalas University Hospital. *Jurnal Fisika Unand*.
  28. Mai, D.V., Drami, I., Pring, E.T., Gould, L.E., Lung, P., Popuri, K., Chow, V., Beg, M.F., Athanasiou, T., & Jenkins, J.T. (2023). A Systematic Review Of Automated Segmentation Of 3d Computed-Tomography Scans For Volumetric Body Composition Analysis. *Journal Of Cachexia, Sarcopenia And Muscle*, 14, 1973–1986.
  29. Massanes, F., & Brankov, J.G. (2016). . Full Receiver Operating Characteristic Curve Estimation Using Two Alternative Forced Choice Studies. *Journal Of Medical Imaging*, 3.
  30. Mehta, R., Kawakita, T.A., & Pineda, A.R. (2024). Modeling Human Observer Detection For Varying Data Acquisition In Undersampled Mri For Two-Alternative Forced Choice (2-AFC) And Forced Localization Tasks. *Medical Imaging*.
  31. Moreau, J., Mechtouff, L., Rousseau, D., Eker, O. F., Berthezene, Y., Cho, T. H., & Frindel, C. (2025). Contrast Quality Control For Segmentation Tasks Based On Deep Learning Models—Application To Stroke Lesions In Ct Imaging. *Frontiers In Neurology*, 16, 1434334.
  32. Palazzo, S., Zambetta, G., & Calbi, R. (2024). An Overview Of Segmentation Techniques For Ct And Mri Images: Clinical Implications And Future Directions In Medical Diagnostics. *Medical Imaging PROCess & Technology*.
  33. Park, S.H., Goo, J.M., & Jo, C. (2004). *Receiver Operating Characteristic (ROC) Curve: Practical Review For Radiologists*. *Korean Journal Of Radiology*, 5, 11 - 18.
  34. Pelberg, R. (2015). Basic Principles In *Computed Tomography (Ct)*. In *Cardiac Ct Angiography Manual* (Pp. 19–58). Springer London.
  35. Romans, L. E. (2010). *Computed Tomography For Technologists*. United Kingdom: Lippincott Williams & Wilkins.
  36. Seeram, E. (2010). *Computed Tomography: Physical Principles And Recent Technical Advances*. *Journal Of Medical Imaging And Radiation Sciences*, 41(2), 87-109.
  37. Seibert, J.A. (2004). X-Ray Imaging Physics For Nuclear Medicine Technologists. Part 1: Basic Principles Of X-Ray Production. *Journal Of Nuclear Medicine Technology*, 32(3), 139-47.
  38. Shao, M., Byrd, D.W., Abbey, C.K., Mitra, J., Wollenweber, S.D., Behnia, F., Lee, J.H., Iravani, A., Sadic, M., Chen, D.L., Kinahan, P.E., & Ahn, S. (2023). Deep Learning Model Observers Trained With Human Observer Data From Two-Alternative Forced Choice (2afc) Trials. *2023 Ieee Nuclear Science Symposium, Medical Imaging Conference And International Symposium On Room-Temperature Semiconductor Detectors (Nss Mic Rtsd)*, 1-1.
  39. So, C. (2020). Aesthetic Preferences Of Neural Style Transfer-Generated Portrait Images: An Exploratory Study With The Two-Alternative-Forced-Choice Task. *International Conference On Innovative Computing And Cloud Computing*.
  40. So, C. (2022). Measuring Aesthetic Preferences Of Neural Style Transfer: More Precision With The Two-Alternative-Forced-Choice Task. *International Journal Of Human-Computer Interaction*, 39, 755 - 775.
  41. Song, Q., & Gong, C. (2024). Image Reconstruction Method For Incomplete Ct Projection Based On Self-Guided Image Filtering. *Medical & Biological Engineering & Computing*.
  42. Springer Handbook Of Medical Technology. (2011). Germany: Springer Berlin Heidelberg
  43. Supardi, J.S. (2014). Classification Of Ct Scan Images Using Artificial Neural Networks.
  44. Traficante, D., Zoccolotti, P., & Marinelli, C.V. (2025). Letter And Word PROCessing In Developmental Dyslexia: Evidence From A Two-Alternative Forced Choice Task. *Children*, 12.

45. Wahyuni, S., & Amalia, L. (2022). Development And Working Principles Of Computed Tomography (Ct scan). *Galenical: Malikussaleh Student Journal Of Medicine And Health*.
46. Wasserthal, J., Breit, H., Meyer, M.T., Pradella, M., Hinck, D., Sauter, A.W., Heye, T., Boll, D.T., Cyriac, J., Yang, S., Bach, M., & Segeroth, M. (2022). Totalsegmentator: Robust Segmentation Of 104 Anatomic Structures In Ct Images. *Radiology. Artificial Intelligence*, 5(5), E230024.
47. Willemink, M.J., Jong, P.A., Leiner, T., Heer, L.M., Nievelstein, R.A., Budde, R.P., & Schilham, A.M. (2013). Iterative Reconstruction Techniques For Computed Tomography Part 1: Technical Principles. *European Radiology*, 23, 1623-1631.
48. Xi, O., Gu, D., Li, X., Zhou, W., Chen, Q., Zhan, Y., Zhou, X.S., Shi, F., Xue, Z., & Shen, D. (2024). Towards A General Computed Tomography Image Segmentation Model For Anatomical Structures And Lesions. *Communications Engineering*, 3. Based On Abdominal Kernels For Ct Image Noise Reduction. In *2019 International Conference Of Artificial Intelligence And Information Technology (Icaait)* (Pp. 293-296). Ieee.
49. Zain, N.J., Anam, C., Suseno, J.E., Amilia, R., & Naufal, A. (2025). The Effect Of Reconstruction Kernel On Low-Contrast Detectability In Acr Ct Phantom Images Evaluated Using The Two-Alternative Forced Choice Method. *International Journal Of Scientific Research In Science And Technology*.
50. Zhang, H. M., & Dong, B. (2023). 5 - Deep Learning For Ct Image Reconstruction From Part Ii - Deep-Learning Architecture For Various Imaging Architectures. *Cambridge University Press Deep Learning For Biomedical Image Reconstruction*, Pp. 89–113.
51. Yulianto, P. (2008). Image Enhancement To Clarify Brain Ct Scan Photos. *Semesta Teknika*, 11(1), 86-95.
52. Yunardi, R. T., Istiqomah, Q., & Latifah, R. (2019, March). Contrast-Enhanced

How to cite this article: Yanwar Budhiyanto, Eko Hidayanto, Catur Edi Widodo, Choirul Anam. Comparison of two alternative forced choice (2-AFC) and receiver operating characteristic (ROC) Methods for low contrast detectability using the Cathpan 503 Phantom. *International Journal of Research and Review*. 2026; 13(3): 408-413. DOI: <https://doi.org/10.52403/ijrr.20260346>

\*\*\*\*\*