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

Task- and Usage-Driven Assessment of Low-Contrast Detectability in Fluoroscopy Using AI-Based Phantom Registration and Human Visibility Probabilities

Elisa Eccher, MSc[1]; Francesca Saltori, MSc[1]; Daniele Ravanelli, MSc[1]; William Avallone, MSc[1]; Annalisa Trianni, MSc[1].
[1]Medical Physics Department, S.Chiara Hospital, ASUIT, Trento, Italy

ABSTRACT

Low-contrast visibility in fluoroscopy quality control is commonly assessed by human scoring or contrast-to-noise ratio (CNR), although a single CNR threshold may not fully represent detectability across different systems and acquisition conditions.

A TOR-FG18 phantom was imaged using multiple fluoroscopy systems and protocols. A YOLO-based model automatically localized phantom structures, enabling template registration and extraction of CNR, local noise, and apparent spatial resolution (MTF50).
Insert visibility was assessed by an experienced medical physicist.

Visible and non-visible inserts showed significantly different CNR distributions (p<0.01), and CNR provided strong discrimination (AUC≈0.94), with an empirical operating point near 0.6. However, overlap persisted at intermediate CNR values.
Logistic regression showed independent contributions from log(CNR) and MTF50, with higher resolution associated with greater detection probability at comparable CNR.

CONTACT

Annalisa Trianni
Medical Physics Department (ASUIT)
S.Chiara Hospital, 38122, Trento, Italy
segreteriafisicasanitaria@apss.tn.it
+39 0461 903282

INTRODUCTION

Low-contrast visibility is an important component of image-quality assessment in fluoroscopy quality control [1]. It is commonly evaluated by visual scoring of contrast-detail phantom inserts. Although this approach is directly related to observer performance, it is time-consuming and may be influenced by acquisition conditions and reader variability.

CNR provides an objective alternative, but it reduces image quality to a single metric and does not explicitly account for spatial resolution. Consequently, a fixed CNR threshold may not fully describe whether a low-contrast object is visible across different fluoroscopy systems and protocols.

Automated phantom localization can improve the reproducibility of quantitative measurements, while human observer labels can be used to relate image-derived metrics to the probability of detection.

Objective: To develop an automated TOR-FG18 analysis workflow and investigate whether a probabilistic approach incorporating multiple image-quality metrics provides a more informative description of low-contrast visibility than a single metric alone.

METHODS AND MATERIALS

Phantom acquisitions: A TOR-18FG image-quality phantom was imaged on multiple C-arm fluoroscopy systems under different acquisition conditions and protocols, generating approximately 140 images and 2520 insert-level observations. The acquisition settings were selected to reflect the variability encountered during routine quality-control measurements and clinical system use.

Automated phantom analysis: A YOLO-based localization model identified predefined reference structures within each phantom image [2].

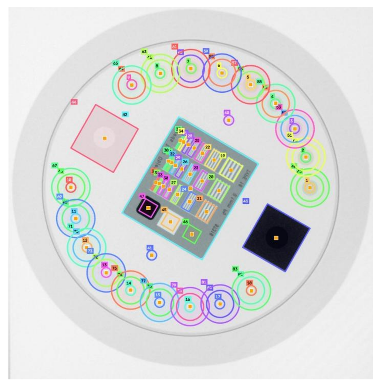

Figure 1. Template mask.

The detected coordinates were used to register a template mask (Figure 1) and automatically propagate regions of interest to the low-contrast inserts, corresponding local-background areas, and high-contrast line-pair targets.

For each low-contrast insert, the automated workflow calculated:

- CNR;
- local background noise;
- apparent spatial resolution, expressed as MTF50 and derived from the line-pair targets using Droege's method.

Insert visibility was independently classified as visible or non-visible by an experienced medical physicist. These binary insert-level labels were used as the reference observer outcome.

Statistical analysis: CNR distributions were compared using the Wilcoxon–Mann–Whitney test. ROC analysis was performed to assess the discriminatory ability of CNR and to identify an empirical operating point using Youden's index. Conditional detection probabilities were evaluated after stratification by local noise and MTF50. Finally, a multivariable logistic regression model was fitted using visibility as the outcome and log(CNR) and MTF50 as predictors. Model coefficients and performance estimates were reported with 95% confidence intervals.

RESULTS

CNR distributions differed significantly between visible and non-visible inserts (Wilcoxon–Mann–Whitney, $p < 0.01$; Figure 2).

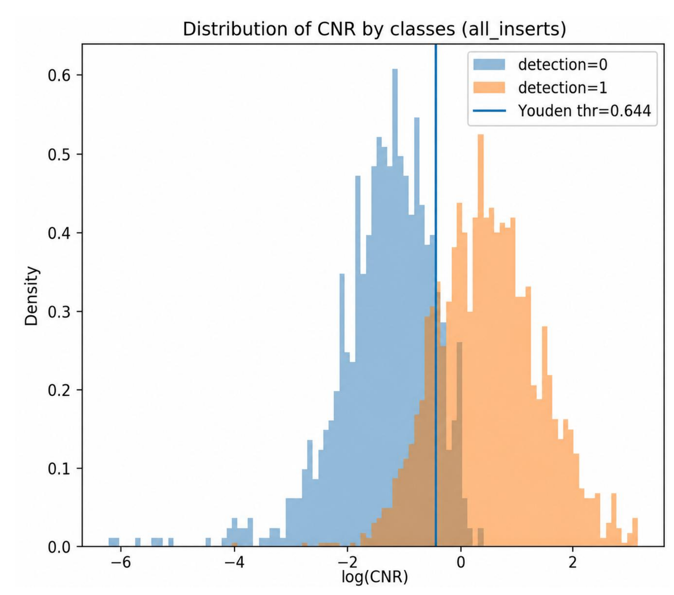

Figure 2. CNR distribution for visible (orange) and non-visible (blue) TOR-18 low-contrast inserts.

Nevertheless, the distributions showed a residual overlap of 24%, increasing to 45% when the analysis was restricted to the nine least visible inserts.

ROC analysis confirmed that CNR was a strong discriminator of visibility across the examined acquisitions, with an apparent AUC of approximately 0.94 (95% CI ≈ 0.94–0.95; Figure 3a).

The Youden operating point was around CNR=0.6, corresponding to a sensitivity of approximately 0.84 and a specificity of approximately 0.89.

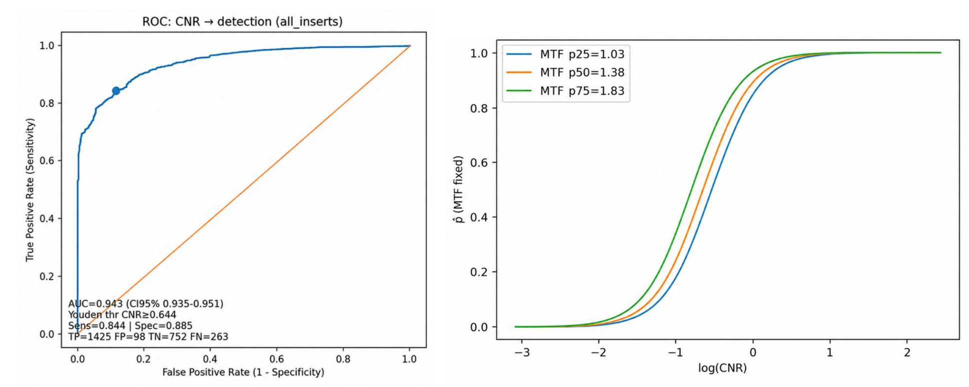

Figure 3. a) ROC analysis (left); **b)** detection probability from logistic regression on log(CNR) for fixed MTF50 values (25th, 50th and 75th percentile of the MTF50 dataset distribution) (right).

Conditional probability analysis indicated that detectability also varied with local noise and apparent spatial resolution, with higher detection probabilities observed in higher-resolution classes under comparable contrast and noise conditions.

In the multivariable logistic model, both log(CNR) and MTF50 contributed independently to visibility (Figure 3b).

DISCUSSION

These findings support a task- and usage-driven interpretation of low-contrast detectability across heterogeneous fluoroscopy systems and acquisition conditions. Although CNR showed strong discriminatory performance, the residual overlap indicates that a single threshold may not adequately represent low-contrast detectability.

Detectability may be better represented as a probability conditioned on multiple image-quality features. The observed association with MTF50 suggests that detectability may reflect image characteristics not fully represented by CNR alone.

AI-based phantom registration supports this approach by improving the consistency of ROI placement and metric extraction across systems and protocols.

However, the probabilistic model remains preliminary. Its interpretation is limited by observer dependence, clustered insert-level observations within the same acquisitions, and the absence of external validation.
Further work is required to assess its calibration, robustness and generalizability across observers, systems, and protocols.

CONCLUSIONS

Automated low-contrast assessment in C-arm fluoroscopy appears feasible. Combining AI-based phantom registration with observer-informed probabilistic modelling provides a reproducible link between quantitative image-quality metrics and human detectability, while accounting for information not fully captured by CNR alone.

REFERENCES

[1] Trianni A, Rogers A, Bertolini M, et al. Quality Control of Dynamic X-Ray Imaging Systems. EFOMP Protocol. Version 01.2024.

[2] Jocher G, Qiu J. Ultralytics YOLO11. Version 11; 2024.