Quantifying the Impact of Energy-Domain Translation in Simulated Digital Mammography

Richard "Ryan" Wargo [1,4], Monica Ghita [1,4], William C. Sleeman IV [3], and Siyong Kim [2,4]

[1] Department of Radiology, School of Medicine, Virginia Commonwealth University, Richmond, VA, USA

[2] Department of Radiation Oncology, School of Medicine, Virginia Commonwealth University, Richmond, VA, USA

[3] Intelligent Health Solutions, LLC, Henrico, VA, USA

[4] Medical Physics Graduate Program, Department of Radiation Oncology, School of Medicine, Virginia Commonwealth University, Richmond, VA, USA

PURPOSE / OBJECTIVES

To quantify the impact of deep learning–based energy-domain translation on image quality and lesion conspicuity in simulated digital mammography using the VICTRE framework.

METHODS

Dataset Generation[1-7]

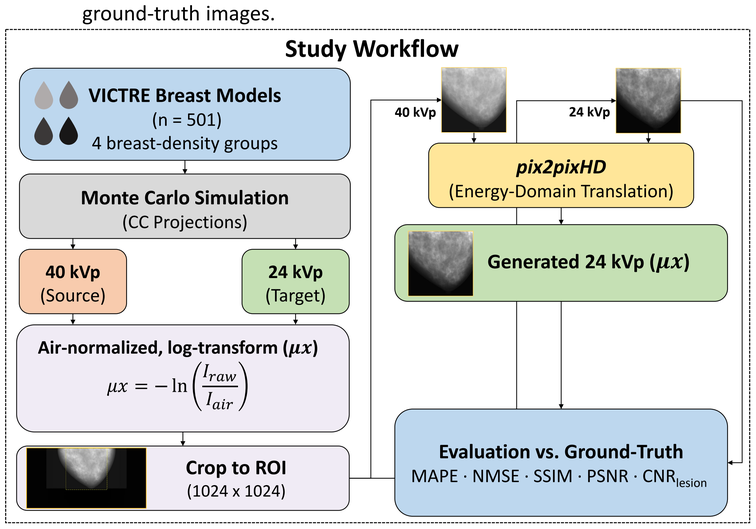
- **Cranio-caudal projection** mammograms generated using **VICTRE framework**
- **501 breast models** spanning four breast-density categories simulated at **24 kVp** and **40 kVp [W/Rh]**
  - Exposure Normalization informed by clinical AEC behavior
- Models divided into training **(n = 451)** and testing **(n = 50)** cohorts
- Lesion-containing subset included **100 mass-lesion models** (70 training, 30 testing)
- Lesion mass densities evaluated at **ρ = 1.06** and **1.25 g/cm³**
- Mean lesion size parameter **(α): 2.32 ± 0.89 mm**

Energy-Domain Translation

- Modified ***pix2pixHD***[8] model trained to translate:
  - Input: 40 kVp **(μx)** projections
  - Output : generated 24 kVp **(μx)** projections
- Projection pairs were cropped to a standardized 1024 × 1024 ROI prior to network training
  - μx -- **Air-normalized, log-transformed attenuation images**
    - $\mu x = -\ln\left(\frac{I_{raw}}{I_{air}}\right)$ where I_{raw} is raw projection image and I_{air} represents the mean detector signal measured in a predefined air-only region
- Model Training:
  - Bayesian optimization used for hyperparameter selection
  - Batch Size = 2; Training Duration = 200 epochs; Initial Learning Rate: $6.1\ x\ 10^{-5}$; Final Learning Rate: $8.8\ x\ 10^{-6}$
- Multiple network configurations were evaluated
  - **Results are reported for the best-performing model (Step 4)**
  - Step 4 incorporated physics-informed μx calibration loss and a breast-masked L1 image loss within the breast region

Quantitative Evaluation

- Translated images were compared with ground-truth 24 kVp projections using:
  - Mean Absolute Percentage Error (**MAPE**), Normalized Mean Squared Error (**NMSE**), Structural Similarity Index (**SSIM**), Peak Signal-to-Noise Ratio (**PSNR**)
  - Lesion preservation was evaluated using **lesion contrast-to-noise ratio** (CNR_{lesion})
- Statistical comparisons were performed between translated and ground-truth images.



RESULTS

Table 1. Best-epoch image-quality metrics for the final energy-domain translation configuration (Step 4) across evaluation datasets.

Model Variant	Dataset	Epoch	MAPE	NMSE	SSIM	PSNR
Step 4	No lesions	185	1.89 ± 0.48	0.0047 ± 0.0017	0.897 ± 0.020	37.49 ± 1.58
	Lesions (d=1.06 g/cm³)	195	2.11 ± 1.28	0.0068 ± 0.0103	0.900 ± 0.021	37.28 ± 2.48
	Lesions (d=1.25 g/cm³)	180	1.95 ± 0.66	0.0052 ± 0.0030	0.901 ± 0.021	37.44 ± 1.84

Notes: Metrics are reported at the epoch corresponding to minimum test-set MAPE for each dataset and were computed within the breast region using a binary breast mask. Consistent performance was observed across lesion-free and lesion-containing datasets.

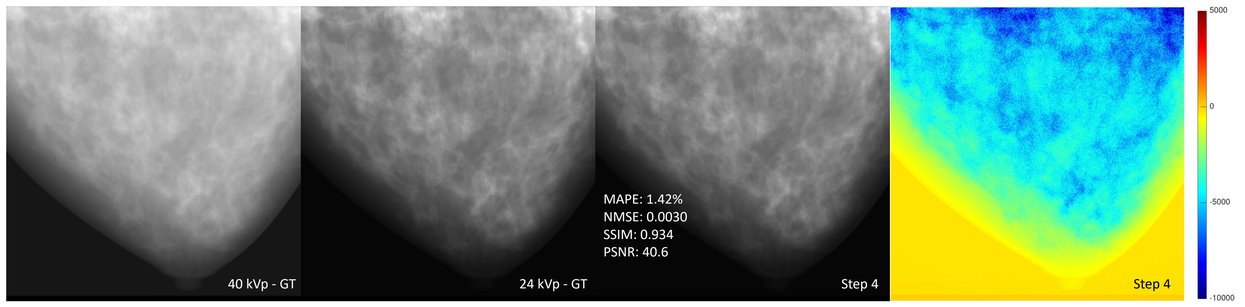


Figure 1. Representative lesion-free energy-domain translation example (seed = 14999). Panels show the 40 kVp source projection, 24 kVp ground-truth projection, Step 4 translated projection, and the corresponding pixel-wise difference map (Δ = GT − model). Image-quality metrics (MAPE, NMSE, SSIM, and PSNR) demonstrate strong agreement between translated and ground-truth images, with residual differences remaining localized relative to the overall image dynamic range. Metrics were computed within a binary breast mask.

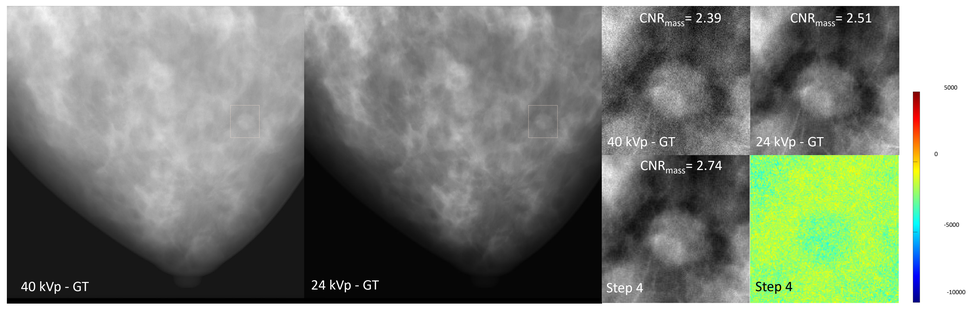


Figure 2. Representative lesion-containing energy-domain translation example (seed = 87902; lesion density = 1.25 g/cm³). Panels show the 40 kVp source projection, 24 kVp ground-truth projection, Step 4 translated projection, magnified lesion ROIs, and corresponding difference maps (Δ = GT − model). Lesion conspicuity and local background structure were preserved following translation, with lesion CNR values comparable to those measured in ground-truth images.

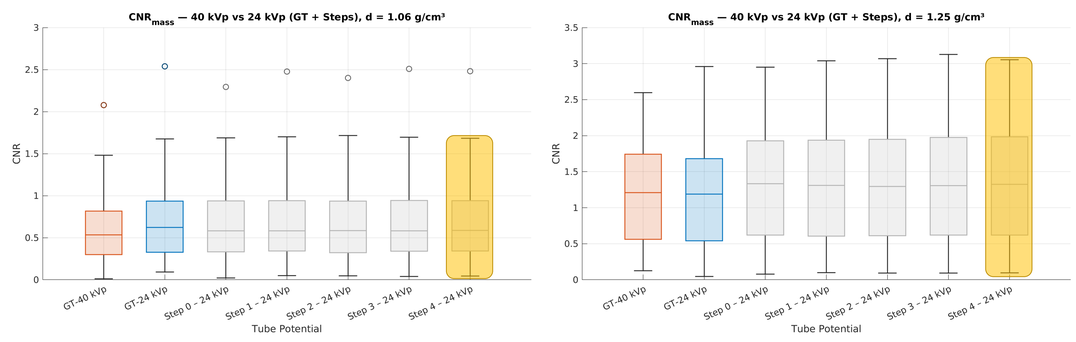


Figure 3. Mass Lesion CNR for translated and ground-truth images. Results are shown for mass lesions with densities of ρ = 1.06 and 1.25 g/cm³. Ground-truth 40 kVp and 24 kVp projections provide reference ranges for interpreting translated 24 kVp outputs. CNR was computed using the lesion-centered background (LC-BG) ROI methodology[9] described in a companion VICTRE lesion-CNR investigation. Lower-density lesions demonstrated no significant degradation in CNR, while higher-density lesions exhibited statistically significant CNR improvements relative to ground truth.

CONCLUSIONS

- **Deep learning–based energy-domain translation accurately reproduced low-energy (24 kVp) mammographic projections** from high-energy (40 kVp) inputs within the VICTRE simulation framework.
- **Global image-quality metrics** demonstrated **strong agreement** with ground truth across lesion-free and lesion-containing datasets **(MAPE ≈ 2%, NMSE ≈ 0.006, SSIM ≈ 0.90, PSNR ≈ 37 dB)**
- Paired statistical comparisons demonstrated **no significant reduction** in **mass-lesion CNR** for **ρ = 1.06 g/cm³ lesions** and **significant CNR increases** for **ρ = 1.25 g/cm³ lesions**.
- These findings establish a **quantitative baseline for task-based evaluation of energy-domain translation in simulated mammography**.

FUTURE WORK

- Evaluate task-based performance using model observers (e.g., CHO) and/or reader studies.
- Expand lesion diversity, particularly microcalcification clusters.
- Extend the framework to additional acquisition geometries and digital breast tomosynthesis (DBT).
- Investigate alternative physics-informed and transformer/diffusion-based architectures for energy-domain translation.

ACKNOWLEDGEMENTS

High Performance Computing resources provided by the HPRC core facility at Virginia Commonwealth University (https://hprc.vcu.edu) were used for conducting the research reported in this work.

REFERENCES

CONTACT INFORMATION

For more information, please contact:

Ryan Wargo, Ph.D., DABR; richard.wargo@vcuhealth.org