GE HealthCare

Clarify DL* –
Deep learning image reconstruction for bone SPECT scans

Omri Ziv, Bella Yuzefovich, Jonathan Sachs, Gil Kovalski

Contents

Abstract

The signal-to-noise ratio (SNR) of SPECT cameras is limited by the system's physics and reconstruction algorithm. Due to the relatively high noise levels in SPECT images, there is a need to use a post-reconstruction mathematical filter to smooth the image and make it more readable. Unfortunately, those filters reduce the image contrast and resolution as well.

In recent years, deep learning techniques have shown great potential in improving the image quality of various medical imaging applications [1-4]. Nevertheless, acquiring high quality SPECT images that will serve as a learning target presents a challenge. To overcome this challenge, GE HealthCare developed a simulations-based deep learning approach designed to improve the image quality of bone SPECT scans - Clarify DL.

This document describes Clarify DL's development and how it is designed to contribute to bone SPECT image quality improvement.

Bone SPECT overview

Bone SPECT, the second most frequently performed SPECT examination in routine nuclear medicine practices, aids in the diagnosis of bone metastasis, infections, inflammation, and fractures [5]. Compared to planar scintigraphy, SPECT increases image contrast and improves lesion detection and localization [6]. Nevertheless, there are ongoing efforts to improve image quality and enable reductions in acquisition times.

Signal and noise tradeoffs in SPECT imaging

Signal-to-noise ratio (SNR) is a quantitative measure used to assess quality of images, including that of nuclear medicine (NM) single photon emission computed tomography (SPECT) images. SNR is the ratio of the strength of the desired signal (i.e., true information of interest) to the level of the background noise (i.e., random variations not related to the signal). A higher SNR indicates that the desired signal is more visible relative to the background noise and is therefore easier to identify. Improvement in SNR can be achieved by increasing the signal and/or by suppressing noise levels.

SNR in SPECT imaging is primarily dictated by the system's design, patient-related factors, injected dose, acquisition time, and by the algorithms used for image reconstruction and processing. The achievable SNR is limited by the statistical photonic noise that is inherently present in SPECT imaging due to the probabilistic nature of photon emission and detection, which is bounded by the system design. The noise is normally reduced to the readers' preferred levels by applying conventional noise reduction techniques, such as 3D post-filters and image reconstruction regularization. These techniques are limited because the noise level may be similar in characteristics to those of small findings (e.g., lesions), thus applying these techniques may reduce noise at the cost of image contrast and resolution. **Figure 1** illustrates the effect of a Butterworth (BW) 3D post-filter on a sample clinical bone SPECT study.

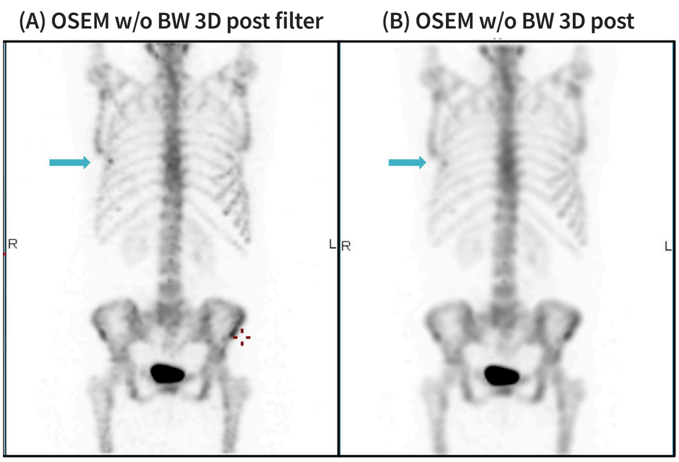


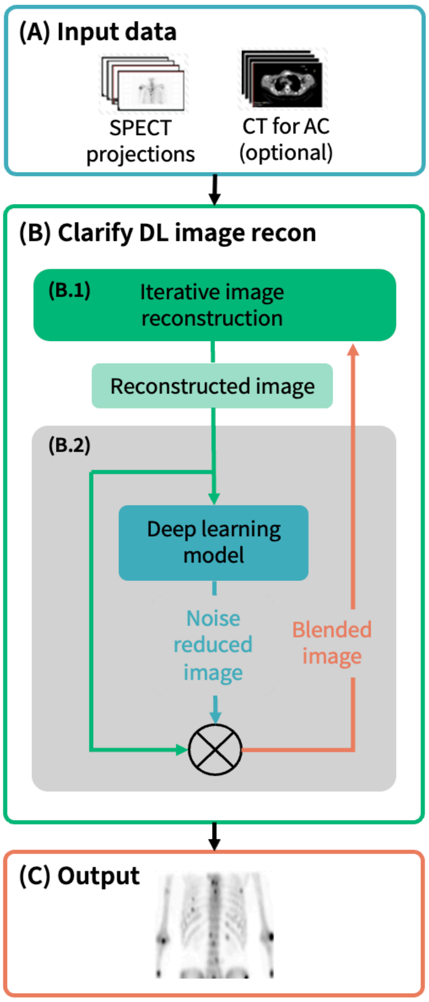
Figure 1: Bone SPECT/CT obtained on NM/CT 870 DR and reconstructed using ordered subset expectation maximization (OSEM) with attenuation correction (AC) and resolution recovery (RR). The blue arrows point to a lesion of impaired contrast and resolution in Image B with BW 3D post-filter, compared to Image A without BW 3D post-filter. Image A is, on the other hand, noisier than Image B

Therefore, because of the limitations with conventional noise reduction techniques, there is a need for a solution that promotes noise reduction without impairing clinical image contrast and resolution.

Clarify DL overview

Clarify DL is designed to be a deep learning-based reconstruction method of NM bone SPECT scintigraphy images obtained using GE HealthCare SPECT and SPECT/CT systems. Clarify DL image reconstruction is made available in addition to the existing reconstruction method*. Clarify DL is designed to reduce image noise while maintaining contrast to enable increased contrast-to-noise ratio (CNR) and contrast recovery coefficient (CRC).

The DL model is integrated into the reconstruction's iterative process and applied to the current reconstructed image at a constant interval in between image updates. The model is designed to generate a noise reduced image that is blended with the input image and integrated back into the iterative process of the reconstruction. **Figure 2** below provides an illustration of Clarify DL image reconstruction design.



(A) Input data: NM bone SPECT scintigraphy projections. The input data can also include CT imaging data for attenuaiton correction (AC) of the SPECT images

(B) Clarify DL image recon: (B.1) Iterative image reconstruction: The SPECT projections are reconstructed using iterative image reconstruction (e.g., OSEM). (B.2) Deep learning model: At a constant update interval, Clarify DL's deep learning model is applied to the reconstructed image to generate a noise reduced image. The noise reduced image is blended with the current reconstructed image to generate a blended image that is used to resume the iterative image reconstruction

(C) Output: SPECT image reconstructed with Clarify DL. The output image is either the last reconstructed image or the blended image, depending on the interval of step (B.2)

Clarify DL design

Clarify DL's image reconstruction model interval and scatter correction parameters can be modified by the user. These two parameters were pre-populated by GE HealthCare to provide optimal image reconstruction; however, users can modify them if desired. These parameters and their effect are summarized in **Table 1**. Other parameters (e.g., number of iterations and subsets) are fixed and cannot be modified.

Parameter	Description
Model interval	The model interval parameter sets the frequency of applying Clarify DL's deep learning model in between image updates. A higher model interval decreases the level of image noise control. A lower model interval increases the level of image noise control.
Scatter correction	The scatter correction parameter determines if scatter correction is to be applied during the iterative image reconstruction. Scatter correction increases quantitation accuracy, if performed, at the cost of subtracting photons.

Table 1: Clarify DL user adjustable parameter

Figure 3 below illustrates the effect of the model interval parameter using a sample clinical case that was reconstructed using three different model interval values.

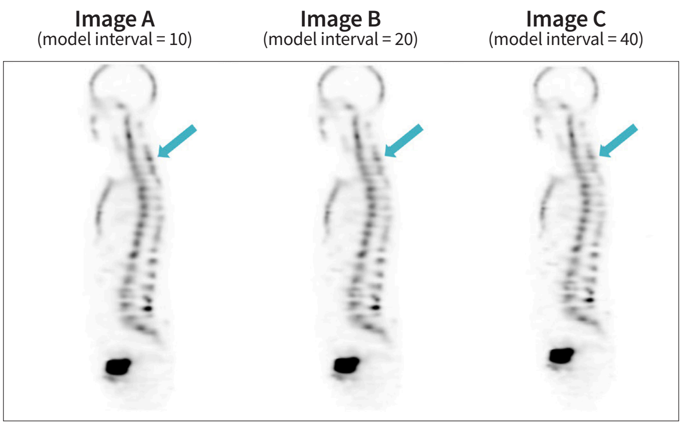


Figure 3: Sample clinical case reconstructed using different model intervals. Images A, B, and C were reconstructed while applying Clarify DL's deep learning model every 10, 20, and 40 image updates. The blue arrow points to an uptake that is sharper on Image A, which is reconstructed with more frequent applications of the model, as compared to Images B and C (i.e., interval of 10 vs. 20 or 40)

Model training method

Clarify DL's deep learning model was trained in supervised learning sessions and designed to reduce noise in NM bone SPECT scintigraphy images. The training used simulated SPECT image reconstructions generated from patients' actual clinical NM bone SPECT scintigraphy studies. The simulated images included pairs of noisy and noise-reduced images that respectively served as the input and target data for the model training (i.e., noisy simulated images serve as input data, while the noise-reduced, high lesion contrast, "clean" simulated images serve as target data).

The use of clinical studies-based simulated images addresses the primary challenge of generating high quality target ground truth images, which have high SNR images (low background noise and high lesion-to-background ratio). This is considered the primary challenge because obtaining such ground truth images requires long duration SPECT acquisitions that are impractical to obtain in clinical settings. This approach also contributes to the model's generalizability because it enables accounting for the primary sources of variations such as lesion contrast-to-background ratio and count density, which simulate acquisition duration or injected dose. These sources of variations can impact the image quality of NM bone SPECT scintigraphy imaging.

Data sets

The patients' clinical studies were divided into three independent data sets, which include a training set, validation set, and testing set. Simulated images generated from the same patient study (i.e., realizations) were only used in the data set to which the patient study was allocated. The patients' actual clinical images were not used in any of the data sets.

Data generation

The process used to generate the simulated SPECT reconstructions for the training and validation sets is described in **Figure 4** below.

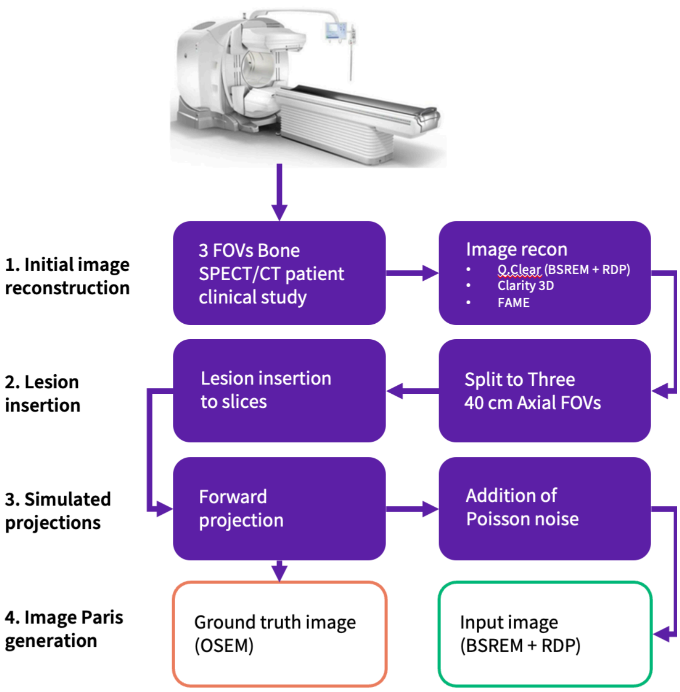


Figure 4: Clarify DL training data generation process

As shown in Figure 4, the simulated SPECT reconstructions are generated in four steps as follows:

1. **Initial image reconstruction:** The patient clinical SPECT/CT exam is reconstructed using block sequential regularized expectation maximization (BSREM) with relative difference prior (RDP) noise regularization (Q.Clear) with post-processing algorithms to generate low-noise, edge-preserved images. For practical purposes, these images can be treated as “noise free images” for the simulation of projections.

2. **Lesion insertion:** Each noise free object is divided into axial field of views (FOV) to enable estimation of a more realistic scan geometry for the forward projection. For each FOV, three simulated lesions were added in regions where lesions are typically found on “clinical” bone scans. The locations, sizes, and shapes of the simulated lesions were defined by a qualified NM physician. The lesion insertion was performed in the image domain by defining the voxels while considering the desired location and size. For each inserted lesion, the lesion to normal bone (L2B) ratio was randomly selected from a set of clinically representative ratios in the range of 1:1 (i.e., no lesion) to 8:1. Ratios above 8:1 were not included because the lesion visibility is not expected to significantly change between different image reconstruction methods.

   For each FOV, a total of five images were generated, each with a different combination of L2B ratios. This resulted in 6 noise free images per FOV, including the image with no inserted lesions.

3. **Simulated projections:** The noise free images were forward projected by using the systems’ collimator and detector characteristics. Each simulated projection was scaled to the count density equivalent to a scan of a predetermined duration based on the original patient clinical scan. Finally, Poisson noise was added to the resulting noise free projections to create an additional set of “noisy” projections.

4. **Image pairs generation:** The simulated projections, with and without noise, were reconstructed to create pairs of input images and target ground truth images to be used during the training. The parameters used are those of the factory preset for clinical use.

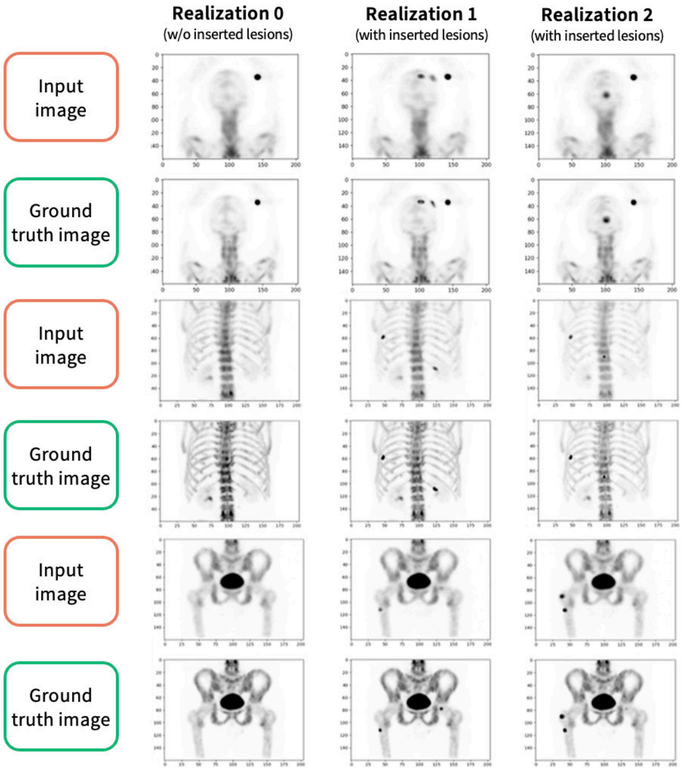


Figure 5: Example of training pair realizations for head, torso, and pelvis. Realization 0 is the image without inserted lesions, while Realizations 1 and 2 are images with inserted lesions and different L2B ratios

A total of 2184 FOVs were generated based on the described process. The data includes simulations of different acquisition durations, lesion combinations, and noise. For the DL model development, the data was divided into three independent data sets: training (1482 FOVs), validation (378 FOVs) and testing (324 FOVs).

DL network

Clarify DL’s model was trained using the 3D residual U-Net convolutional encoder-decoder network. The network is composed of convolution (Conv) layers, batch normalization (BN) and rectified linear unit (ReLU) activation functions, max pooling, up-sampling layers, skip connections, and 4 scales (i.e., 4 max pooling layers). The input images are fed into the network to predict a residual image, which is added to the input. The training of the Clarify DL model is supervised. The output of the network was compared to the target images based on mean squared error (MSE) loss function and was used to update the network’s trainable parameters by backpropagation, as shown in **Figure 6** below.

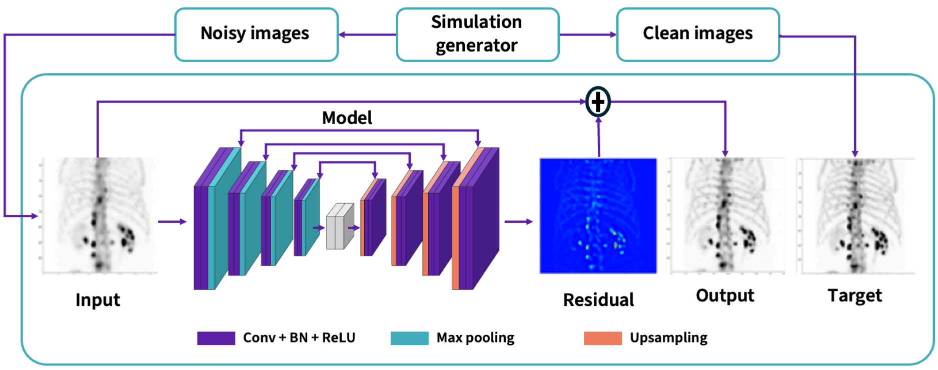


Figure 6: Clarify DL supervised learning training process

Model cross-validation

A 5-folds cross validation method was used to assess the model's performance and generalizability, and to help detect potential overfitting or underfitting by estimating the model's performance across different samples.

Inferencing workflow

Clarify DL's inferencing workflow consisted of 3 steps including pre-inferencing, inferencing, and post-inferencing. During the pre-inferencing step, the reconstruction process is running until the Clarify DL condition is applied. If the updates counter is equal to the model interval parameter, the temporal image is saved to initiate the inferencing step. During the inferencing step, the input image from the pre-inferencing step is processed and prepared for Clarify DL's model, inferenced by the model, and reprocessed back to its original dimensions. The Clarify DL pretrained model is imported. During the post-inferencing step, the reconstruction process resumes with the image generated using Clarify DL's model. To avoid fast convergence of the reconstruction algorithm (mainly with OSEM), the generated image is blended with the input image generated in the pre-inferencing to maintain some of the image noise.

The blended image is sent back to the image reconstruction process as a starting point for the next iteration. When the model interval is such that Clarify DL runs at the end of the image reconstruction process, the output of the inferencing step is the final reconstructed image.

Testing strategy

The performance testing for the final Clarify DL model includes bench and clinical testing. The performance testing used entirely new data and tests that were not part of the training, validation, and test sets described above. The bench and clinical testing included data and tests applicable for dual-head (NaI and CZT) and 3D ring-design (CZT) GE HealthCare scanners, and to both AC and NC imaging data.

Bench testing

Testing used a digital phantom with inserted lesions [**Figure 7**] to confirm improvements over the existing factory presets using cleared image reconstruction methods. The bench performance testing included IQ performance measurements at both the image and lesion levels for AC and NC images. The image level measurements included MSE, structural similarity index measure (SSIM), and peak signal to noise ratio (PSNR) [7]. The lesion level measurements included CNR and CRC of lesions regions of interest (ROI) and their pair background ROIs.

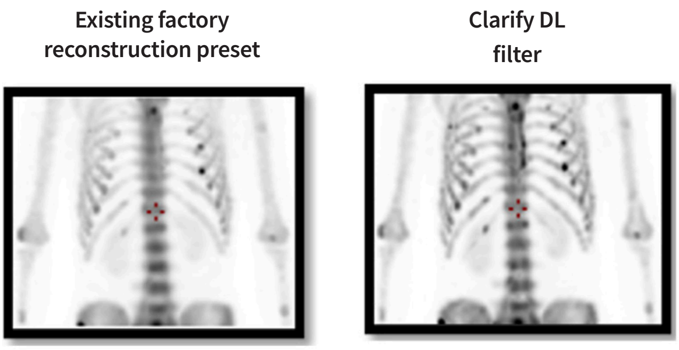


Figure 7: Digital phantom images reconstructed with existing factory reconstruction preset (left) and with Clarify DL reconstruction (right)

Parameter	NCRR vs NCRRDL (% improvement)	ACRR vs ACRRDL (% improvement)
Image SSIM	13.8%	7.9%
Image MSE	72.2%	53.9%
Image PSNR	14.4%	16.8%
Lesions CNR	67.2%	28.0%
Lesions CRC	162.6%	56.2%

Table 2: A comparison of image quality and lesion measurements – SSIM, MSE, PSNR, CNR and CRC

The tests results show that Clarify DL achieves improvement for all obtained image level and lesion level measurements.

Clinical testing (Reader study)

The overall image quality of Clarify DL was reviewed and evaluated by certified NM physicians. The physicians answered blinded preference questions comparing clinical cases of the same studies reconstructed with the factory presets reconstruction (reference) and with Clarify DL.

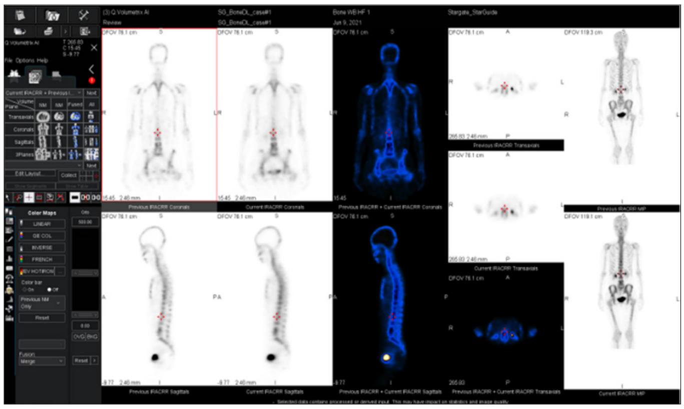


Figure 8: An example of blind comparison review mode for the Clarify DL evaluation process

Clarify DL – Deep learning image reconstruction for bone SPECT scans

127 bone SPECT exams were scored by 9 readers (each exam was scored by 3 readers, each reader read 40-47 exams) for image quality, image resolution and noise level, using a 1-5 Likert scoring (1 – lowest, 5 – highest). Each exam received 36 scores (3 readers x 3 image categories x 4 reconstructions). However, if a CT scan was not acquired, the exam received only 18 scores (3x3x2).

In total, 723 scores for non-corrected and attenuation-corrected reconstructions were provided across all parameters of image quality, image resolution, and noise level per system.

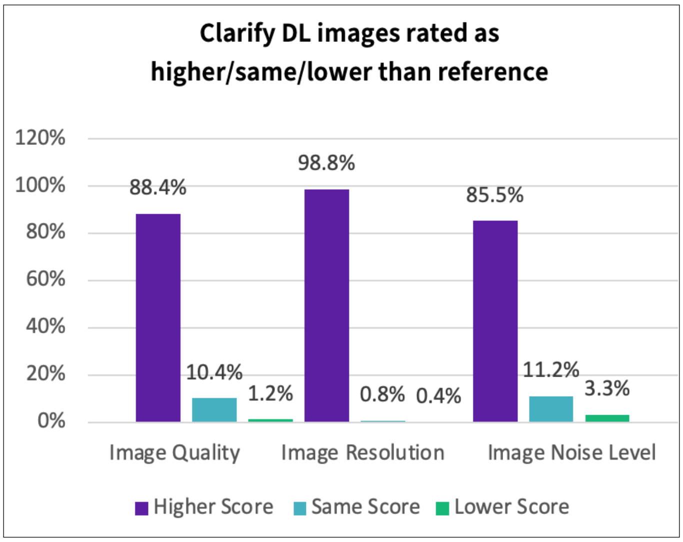


Figure 9: Clarify DL images rated as higher/same/lower than REF

As illustrated in **Figure 9**, the reader evaluation of images reconstructed with Clarify DL demonstrated that images generated by the Clarify DL reconstruction algorithm were of diagnostic image quality. Moreover, there was a preference among the evaluators for the Clarify DL images in terms of overall image quality, resolution, and noise level.

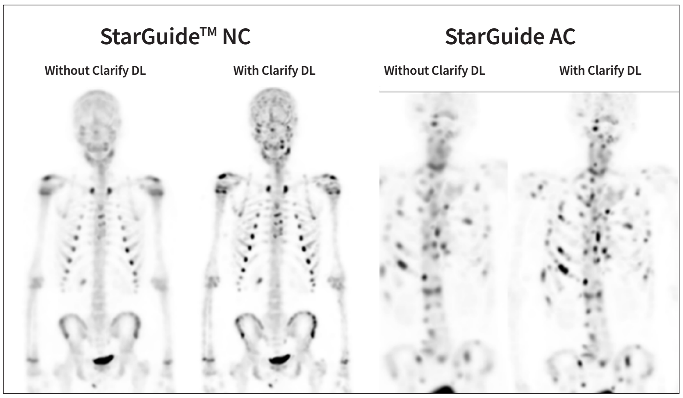


Figure 10: An example of the impact of Clarify DL reconstruction on clinical data. Left pair: Non attenuation correction scans without and with Clarify DL. Right pair: Attenuation corrected scans without and with Clarify DL.

Quantitation

To evaluate Clarify DL behavior for quantitation purposes, recovery coefficient activity measurements using phantom scans (NEMA IEC PET body phantom and Jaszczak phantoms [8], [9], Data Spectrum, Durham, NC, USA) were performed with and without Clarify DL for StarGuide and dual-head systems.

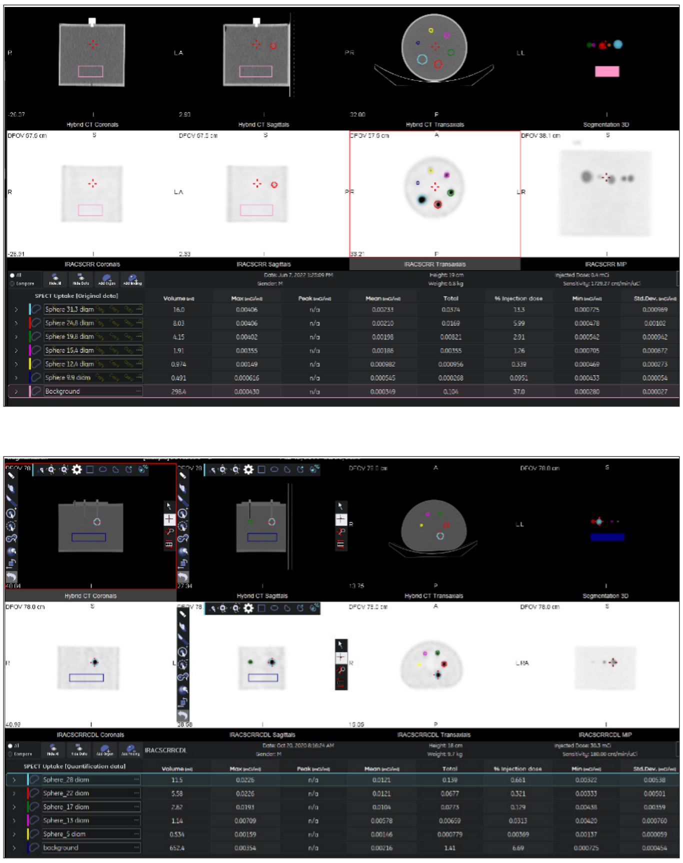


Figure 11: Recovery coefficient measurements on Jaszczak (up) and IEC (down) phantoms

	Sphere diameter (mm)	Volume (ml)	% CRC Diff Clarify DL - Ref
Sph_JZK_60m	31.3	16.06	13.61
Sph_JZK_60m	24.8	7.99	19.73
Sph_JZK_60m	19.8	4.06	29.25
Sph_JZK_60m	15.4	1.91	30.12
Sph_JZK_60m	12.4	0.99	12.48
Sph_JZK_60m	9.9	0.51	2.28
BG		298.4	4.57

Table 3: StarGuide quantitation accuracy comparison between Clarify DL and factory preset reconstruction (reference)

	Sphere diameter (mm)	Volume (ml)	% CRC Diff Clarify DL - Ref	
			NaI	CZT
Sph_28	28	11.5	6.54	9.32
Sph_22	22	5.58	11.3	15.84
Sph_17	17	2.62	10.76	11.65
Sph_13	13	1.14	0.35	4.66
Sph_10	10	0.534	-0.42	0.37
BG		653	0.46	1.09

Table 4: Dual-head NaI and CZT quantitation accuracy comparison between Clarify DL and baseline reconstruction

As illustrated in **Table 3** and **Table 4**, the results show improvement in quantitation accuracy in 18 of the 19 measurements when phantom data was reconstructed using Clarify DL for both dual-head cameras (NaI and CZT) and for 3D ring-design cameras (StarGuide).

Application and future directions

Clarify DL model architecture and training is a general approach which may be applied to other isotopes and clinical indications of SPECT/CT and PET/CT acquisitions.

Conclusion

Clarify DL is a deep learning-based image reconstruction algorithm designed for image quality enhancement of bone SPECT and SPECT/CT studies. The evidence has demonstrated that Clarify DL improves different aspects of image quality: structure, lesion contrast and quantitation in the physics test, and noise and resolution in the clinical reader studies. Improvement was achieved across different SPECT/CT cameras for both AC and NC scans.

References

[1] Mehranian, A.; Wollenweber, S.D.; Walker, M.D.; et al. "Image enhancement of whole-body oncology [18F]-FDG PET scans using deep neural networks to reduce noise." Eur J Nucl Med Mol Imaging 49, 539–549 (2022)

[2] Mehranian, A.; Wollenweber, S.D.; Walker, M.D.; et al. "Deep learning–based time-of-flight (ToF) image enhancement of non-ToF PET scans." Eur J Nucl Med Mol Imaging 49, 3740–3749 (2022)

[3] Tian, Qian; et al. "Image quality improvement in low-dose chest CT with deep learning image reconstruction." Journal of Applied Clinical Medical Physics 23.12 (2022)

[4] Chen, Z.; Pawar, K.; Ekanayake, M.; et al. "Deep Learning for Image Enhancement and Correction in Magnetic Resonance Imaging—State-of-the-Art and Challenges." J Digit Imaging 36, 204–230 (2023)

[5] Bahk, Yong-Whee. "Combined scintigraphic and radiographic diagnosis of bone and joint diseases: including gamma correction interpretation." Springer Science & Business Media (2012)

[6] Sarikaya, Ismet; Ali Sarikaya; Holder; Lawrence E. "The role of single photon emission computed tomography in bone imaging." Seminars in nuclear medicine. vol. 31. no. 1. WB Saunders (2001)

[7] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. "Image quality assessment: from error visibility to structural similarity." IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612 (April 2004)

[8] https://www.spect.com/pdf/NEMA-IEC-PET-Body-Phantom.pdf

[9] https://www.spect.com/pdf/Flanged-Jaszczak-Phantoms.pdf

Acronyms list

AC – Attenuation Corrected

BN – Batch Normalization

BW – Butterworth

CNR – Contrast to Noise Ratio

CRC - Contrast Recovery Coefficient

CT - Computed Tomography

CZT - Cadmium Zinc Telluride

DL – Deep Learning

FOV – Field of View

IQ – Image Quality

L2B – Lesion to Background

MSE - Mean Squared Error

NaI – Sodium Iodide

NC – Non-Corrected

NM – Nuclear Medicine

OSEM – Ordered Subset Expectation Maximization

PSNR - Peak Signal to Noise Ratio

ReLU - Rectified Linear Unit

ROI – Region of Interest

RR – Resolution Recovery

SNR – Signal to Noise Ratio

SPECT – Single Photon Emission Computed Tomography

SSIM - Structural Similarity Index Measure

* Clarify DL may not be available for sale in all regions.

Clarify DL is CE marked with StarGuide and Xeleris V. Clarify DL is not approved or cleared by the U.S. FDA and is not licensed in accordance with the Canadian law. Not available for sale in the United States and Canada.

JB29096XX