Improvement of Image Quality in Low-Count Bone Scintigraphy Using Deep Learning

Taisuke Murata
 Chiba University Hospital

Masahisa Onoguchi (✉ onoguchi@staff.kanazawa-u.ac.jp)
 Department of Quantum Medical Technology, Institute of Medical, Pharmaceutical and Health Sciences,
Kanazawa University   https://orcid.org/0000-0001-7508-5975

Takuma Hashimoto
 Chiba University Hospital

Takayuki Shibutani
 Kanazawa University

Takashi Iimori
 Chiba University Hospital

Koichi Sawada
 Chiba University Hospital

Tetsuro Umezawa
 Chiba University Hospital

Yoshitada Masuda
 Chiba University Hospital

Takashi Uno
 Department of Diagnostic Radiology and Radiation Oncology, Graduate School of Medicine, Chiba
University

Abstract

Objective

To improve image quality for low-count bone scintigraphy whole-body images using deep learning and evaluate their applicability in clinical practice.

Methods

Five hundred fifty patients were included in the study. Low-count Original images (75%, 50%, 25%, 10%, and 5% counts) were generated from Reference images (100% counts) using Poisson resampling. Patients were randomly divided into training (500) and evaluation (50) groups. Output (DL-filtered) images were obtained after training with U-Net using Reference images as teacher data. Gaussian-filtered images were generated for comparison. Peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) to the Reference image were calculated to determine image quality. Artificial neural network (ANN) value, bone scan index (BSI), and number of hotspots (Hs) were computed using BONENAVI analysis for patients with and without bone metastases, to assess diagnostic performance. Accuracy of bone metastasis detection and area under the curve (AUC) were calculated. Original, Gaussian-filtered, and DL-filtered images were compared with Reference images.

Results

PSNR and SSIM for DL-filtered images were highest in all count percentages. BONENAVI analysis values for DL-filtered images did not differ significantly regardless of the presence or absence of bone metastases. BONENAVI analysis values for Original and Gaussian-filtered images differed significantly at < 25% counts in patients without bone metastases. In patients with bone metastases, BSI and Hs for Original and Gaussian-filtered images differed significantly at < 10% counts, whereas ANN values did not. Accuracy of bone metastasis detection was highest for DL-filtered images in all count percentages; AUC did not differ significantly. Accuracy of Original and Gaussian-filtered images worsened with decreasing count percentage; AUC differed significantly for Original images at < 50% counts and for Gaussian-filtered images at < 25% counts.

Conclusions

Our deep learning model improved image quality and bone metastasis detection accuracy for low-count bone scintigraphy whole-body images, suggesting its applicability in clinical practice.

1. Introduction

Bone scintigraphy evaluates various bone lesions, including tumors, arthritis, metabolic bone disease, infection, and trauma, and is frequently used to diagnose metastatic bone tumors [1–4]. However, patients with pain due to bone metastases often cannot tolerate prolonged immobility. The usefulness of single-photon emission computerized tomography (SPECT) imaging has been reported, and examination

throughput improvement is required to incorporate SPECT imaging into the diagnosis pipeline [5–7]. Furthermore, there is concern regarding increased radiation dose, as examinations for follow-up are performed periodically. In nuclear medicine, examination time requires shortening and the amount of radioactivity administered must be reduced, and both factors have the same physical essence regarding their contribution to image quality [8]. However, these factors increase image noise due to a decreased number of acquired counts, resulting in a deterioration of diagnostic performance, owing to image quality degradation.

Smoothing filters used to reduce image noise have the tradeoff of degrading resolution. Additionally, many studies have reported improving nuclear medicine image quality using deep learning [9–15]. These reports mainly used tomographic imaging techniques, such as SPECT and positron emission tomography. For planar images, Ito et al. [16] reported improved image quality of low-count images for thoracic static images in bone scintigraphy. Whole-body imaging is the most valuable imaging technique in bone scintigraphy because it depicts bone lesions throughout the body in a single image. However, as whole-body imaging takes longer than static imaging, it is necessary to reduce the imaging time. The clinical usefulness of computer-aided diagnosis (CAD) software used for diagnosing and following bone metastases has been reported, further heightening the value of whole-body images [17–19].

The present study aimed to develop a deep learning model to improve the image quality of low-count bone scintigraphy whole-body images and evaluate its clinical applicability. We aimed to improve the model's accuracy using clinical images, rather than phantoms, for training.

2. Materials And Methods

The Ethics Review Committee of Chiba University Hospital approved the present retrospective study. All imaging data used for analysis were necessary for clinical diagnosis, and no examinations were performed for this study. The Ethics Review Committee waived written consent.

Patients

A total of 550 patients (49 males and 501 females; median age, 63 years; range, 33–88 years) on whom bone scintigraphy was performed for clinical diagnosis were included. Of these patients, 500 and 50 were randomly selected for training and obtaining evaluation data, respectively. The breakdown of patient diseases in training and evaluation data is shown in Table 1. The diagnosis of the presence of bone metastases was determined by a radiologist using bone scintigraphy, computed tomography, magnetic resonance imaging, blood tests, and clinical findings.

Table 1
Breakdown of patient diseases in training and evaluation data

Disease	Training data			Evaluation data		
	with bone metastasis	without bone metastasis	Total	with bone metastasis	without bone metastasis	Total
Breast cancer	43	383	426	10	32	42
Lung cancer	9	40	49	3	2	5
Prostate cancer	10	4	14	1	0	1
Thyroid cancer	–	1	1	1	–	1
Rectum cancer	2	2	4	–	–	–
Hepatocellular carcinoma	2	2	4	1	–	1
Appendiceal cancer	1	–	1	–	–	–
Plasmacytoma	1	–	1	–	–	–
Total	68	432	500	16	34	50

Data acquisition

All patients were examined 3−4 hours after receiving 699−785 MBq of 99mTc-MDP injection solution (PDRadiopharma Inc., Tokyo, Japan). Anterior and posterior whole-body imaging was performed using a gamma camera (NM/CT 870 DR hybrid SPECT/CT scanner; GE Healthcare, Chicago, IL) with a low-energy high-resolution-sensitivity collimator. Whole-body imaging was performed with a matrix size of 1024 x 256, pixel size of 2.21 mm, 15% energy window centered at the photopeak energy (140.5 keV), and bed speed of 13.3 cm/min.

Low-count image preparation

Low-count Original images were created from all Reference images in patient data (100% counts) using the Poisson resampling method [20] application installed in Xeleris 4DR (GE Healthcare, Chicago, IL). The Poisson resampling method arbitrarily subtracts image counts and adds Poisson noise corresponding to the number of counts. Original images with 75%, 50%, 25%, 10%, and 5% counts were created per patient.

Network architecture and training

We have developed a deep learning model based on U-Net [21]. U-Net was developed for image segmentation; however, it also has models that were developed for noise reduction [22−24]. The structure of our deep learning model is shown in Fig. 1. In our model, unsharp masking [25−28] was incorporated before the final output. The unsharp masking sharpens the Output image by adjusting the difference

between the Output' image and the smoothed Output' image and adding it to the Output' image. The formula for unsharp masking is shown below.

$$f_{uij} = d_{ij} + \delta \sum_{l=-p}^{p} \sum_{m=-p}^{p} w_{lm} \left(d_{ij} - d_{i+l,j+m} \right)$$

Where f_{uij} is the pixel value of the Output image at position (i, j), d_{ij} is the pixel value of the Output' image at position (i, j), δ is a parameter to adjust the degree of sharpening, and w_{lm} is a Gaussian filter; δ was set to 15. Mean square error (MSE) was used as the loss function for learning. ReLU was the activation function. Adam was the optimizer, and the learning rate was 1.0×10^{-6}. The batch size was 64, and the number of epochs was 1000.

Data analysis

Gaussian-filtered images were obtained by applying a Gaussian filter to the Original image at each count percentage. The size of the Gaussian filter was set to 7 mm, as this setting has been reported to have the best correlation with the Reference image [29]. The present study compared the results of the Original image, Gaussian-filtered image, and Deep learning-filtered (DL-filtered) image, i.e., our model's output image.

First, peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) with the Reference image were calculated for the Original, Gaussian-filtered, and DL-filtered images.

PSNR and SSIM were calculated using the following equations.

$$PSNR = 10\log_{10} \left(\frac{I^2}{MSE} \right) \text{[db]}$$

Where I is the maximum count of Reference images.

$$SSIM\left(x,y\right) = \frac{\left(2u_x u_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)}$$

Where x is the Reference image, y is the target image for comparison, μ_x and μ_y are the mean pixel values of x and y, respectively, and σ_x and σ_y are the standard deviations of the pixel values of x and y, respectively. σ_{xy} is the covariance of the pixel values of x and y. C_1 and C_2 are constants, expressed as $C_1 = (K_1 I)^2$ and $C_2 = (K_2 I)^2$, with K_1 and K_2 set to 0.01 and 0.03, respectively, to avoid division due to minimal values. PSNR and SSIM were compared among Original, Gaussian-filtered, and DL-filtered images.

Furthermore, Reference, Original, Gaussian-filtered, and DL-filtered images were analyzed using BONENAVI software (PDR Pharma, Tokyo). The bone segments analyzed were whole-body bones (skull, cervical spine, thoracic spine, lumbar spine, humerus, thorax, pelvic bone, and femur), except for bones of

the peripheral limbs. The artificial neural network (ANN) value, bone scan index (BSI), and hot spot number (Hs) were calculated. The minimum and maximum ANN values were calculated as 0 and 1, respectively; ANN value indicates the confidence level of bone metastasis with a threshold value of 0.5 based on factors such as the shape, location, and count of high accumulated areas [17, 19]. The percentage of bone segment areas with high-risk bone metastasis accumulation sites are shown as BSI and the number as Hs. The analysis values of the Reference image were compared with those of the Original, Gaussian-filtered, and DL-filtered images.

The sensitivity, specificity, and diagnostic accuracy were calculated based on ANN value of > 0.5 indicating the presence of bone metastasis. In addition, receiver operating characteristic (ROC) analysis was performed, and the area under the curve (AUC) was compared.

Statistical analysis

PSNR and SSIM with the Reference image were evaluated using the Steel-Dwass test. ANN values, BSI, and Hs were evaluated using the Steel test, via comparison with those of the Reference image as the control. These values were statistically analyzed using JMP Pro (version 16.1.0; SAS Institute, Cary, NC). DeLong's test was used to evaluate the AUC calculated using ROC analysis. Statistical analyses were performed using EZR (Saitama Medical School Hospital, Saitama, Japan), a GUI of R (The R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was set at $p < 0.05$ for all analyses.

3. Results

Figures 2 and 3 show an example of each image's output for a patient without bone metastasis and a patient with bone metastasis, respectively. The Original image's noise increased as the count percentage decreased. The Gaussian-filtered image suppressed the noise; however, the 10% and 5% count images also suppressed the bone accumulation signal. The DL-filtered images restored image counts, suppressed noise, and preserved bone morphology. The presence of bone metastases in patients with bone metastases could be recognized even in the 5% count images.

The PSNR results are shown in Table 2. The PSNR of the Original and Gaussian-filtered images reduced as the count percentage decreased but was relatively preserved for the DL-filtered images. PSNR of the Gaussian-filtered images was lower than that of the Original images at any count percentage, with significant differences at the 75% and 50% counts. The PSNR of DL-filtered images was significantly higher than that of Original and Gaussian-filtered images at any count percentage.

Table 2
PSNR of Original, Gaussian-filtered, and DL-filtered images as a reference to the Reference image.

Count percentage	Mean ± SD			P-value		
	Original	Gaussian-filtered	DL-filtered	Original vs. Gaussian-filtered	Gaussian-filtered vs. DL-filtered	Original vs. DL-filtered
75%	79.14 ± 3.19	77.2 ± 2.81	85.03 ± 2.56	**< 0.01**	**< 0.01**	**< 0.01**
50%	73.41 ± 3.29	72.77 ± 3.14	80.00 ± 5.23	**< 0.01**	**< 0.01**	**< 0.01**
25%	69.99 ± 3.32	69.69 ± 3.24	80.70 ± 4.36	0.35	**< 0.01**	**< 0.01**
10%	68.44 ± 3.33	68.24 ± 3.27	79.27 ± 3.68	0.50	**< 0.01**	**< 0.01**
5%	67.98 ± 3.33	67.81 ± 3.28	80.44 ± 3.75	0.56	**< 0.01**	**< 0.01**
P-values denoting significant differences are shown in bold.						

The results of SSIM are shown in Table 3. SSIM of the Gaussian-filtered images was lower than that of the Original images at any count percentage, with a significant difference at 75%. The SSIM of DL-filtered images was significantly higher than that of Original and Gaussian-filtered images at any count percentage.

Table 3
SSIM of Original, Gaussian-filtered, and DL-filtered images as a reference to the Reference image.

Count percentage	Mean ± SD			P-value		
	Original	Gaussian-filtered	DL-filtered	Original vs. Gaussian-filtered	Gaussian-filtered vs. DL-filtered	Original vs. DL-filtered
75%	0.99997 ± 0.00003	0.99996 ± 0.00003	0.99999 ± 0.00001	**< 0.01**	**< 0.01**	**< 0.01**
50%	0.99987 ± 0.00014	0.99986 ± 0.00014	0.99997 ± 0.00006	0.19	**< 0.01**	**< 0.01**
25%	0.99971 ± 0.00033	0.99969 ± 0.00033	0.99996 ± 0.00021	0.42	**< 0.01**	**< 0.01**
10%	0.99957 ± 0.00048	0.99956 ± 0.00048	0.99995 ± 0.00030	0.51	**< 0.01**	**< 0.01**
5%	0.99952 ± 0.00053	0.99951 ± 0.00054	0.99996 ± 0.00029	0.56	**< 0.01**	**< 0.01**
P-values denoting significant differences are shown in bold.						

Table 4 shows ANN values, BSI, and Hs results in patients without bone metastases. Original and Gaussian-filtered images overestimated all analytical values as the count percentage decreased; in particular, there was a significant difference with the analytical value of Reference images at < 25% of counts. DL-filtered images tended to overestimate all analytical values as the count percentage decreased; however, there were no significant differences compared to the analytical value of Reference images.

Table 4

Comparison with Reference images of ANN values, BSI, and Hs in patients without bone metastases.

Count percentage	Image type	Mean ± SD			P-value		
		ANN	BSI	Hs	ANN	BSI	Hs
100%	Reference	0.22 ± 0.24	0.18 ± 0.32	2.47 ± 4.78	–	–	–
75%	Original	0.27 ± 0.29	0.20 ± 0.40	2.74 ± 5.43	0.95	0.99	> 0.99
	Gaussian-filtered	0.27 ± 0.32	0.20 ± 0.34	2.21 ± 3.62	> 0.99	0.98	0.86
	DL-filtered	0.20 ± 0.25	0.28 ± 0.56	3.24 ± 6.44	0.88	> 0.99	> 0.99
50%	Original	0.31 ± 0.34	0.30 ± 0.51	4.00 ± 6.13	0.79	0.74	0.34
	Gaussian-filtered	0.34 ± 0.33	0.29 ± 0.46	2.91 ± 4.44	0.54	0.83	> 0.99
	DL-filtered	0.21 ± 0.27	0.37 ± 0.71	3.50 ± 5.34	0.90	0.88	0.8
25%	Original	0.60 ± 0.28	0.57 ± 0.75	10.24 ± 10.99	**< 0.01**	**< 0.01**	**< 0.01**
	Gaussian-filtered	0.56 ± 0.32	0.69 ± 0.96	7.68 ± 10.61	**< 0.01**	**< 0.01**	**< 0.01**
	DL-filtered	0.25 ± 0.27	0.41 ± 0.81	3.94 ± 6.44	0.96	0.47	0.72
10%	Original	0.79 ± 0.15	2.45 ± 1.65	47.5 ± 26.07	**< 0.01**	**< 0.01**	**< 0.01**
	Gaussian-filtered	0.88 ± 0.15	4.63 ± 4.77	51.21 ± 42.24	**< 0.01**	**< 0.01**	**< 0.01**
	DL-filtered	0.26 ± 0.29	0.40 ± 0.76	3.32 ± 5.00	0.99	0.94	0.99
5%	Original	0.84 ± 0.21	4.05 ± 2.85	76.71 ± 51.17	**< 0.01**	**< 0.01**	**< 0.01**
	Gaussian-filtered	0.90 ± 0.16	15.49 ± 8.93	103.21 ± 51.84	**< 0.01**	**< 0.01**	**< 0.01**
	DL-filtered	0.31 ± 0.32	0.33 ± 0.78	2.38 ± 3.82	0.82	> 0.99	0.98

P-values denoting significant differences are shown in bold.

The results of ANN values, BSI, and Hs in patients with bone metastases are shown in Table 5. The ANN values of the Original, Gaussian-filtered, and DL-filtered images were not substantially different from those of the Reference image at any count percentage. For < 10% counts, the BSI and Hs of the Original and Gaussian-filtered images were overestimated and significantly different from the Reference image. BSI and Hs of DL-filtered images were not substantially different from Reference images at any count percentage.

Table 5
Comparison with Reference images of ANN values, BSI, and Hs in patients with bone metastases.

Count percentage	Image type	Mean ± SD			P-value		
		ANN	BSI	Hs	ANN	BSI	Hs
100%	Reference	0.91 ± 0.11	1.37 ± 1.58	11.25 ± 13.63	–	–	–
75%	Original	0.91 ± 0.11	1.34 ± 1.58	12.25 ± 14.69	> 0.99	> 0.99	> 0.99
	Gaussian-filtered	0.95 ± 0.07	1.21 ± 1.34	9.44 ± 9.45	0.71	> 0.99	0.99
	DL-filtered	0.91 ± 0.10	1.53 ± 1.62	12.93 ± 16.15	> 0.99	0.97	> 0.99
50%	Original	0.82 ± 0.23	1.34 ± 1.57	12.31 ± 13.74	0.97	> 0.99	0.94
	Gaussian-filtered	0.95 ± 0.07	1.46 ± 1.36	10.75 ± 10.76	0.47	0.80	> 0.99
	DL-filtered	0.93 ± 0.09	1.79 ± 1.65	15.31 ± 16.65	0.93	0.60	0.80
25%	Original	0.88 ± 0.23	1.54 ± 1.63	17.63 ± 17.98	0.95	0.80	0.33
	Gaussian-filtered	0.95 ± 0.07	1.95 ± 1.77	14.93 ± 14.03	0.37	0.55	0.85
	DL-filtered	0.93 ± 0.10	1.39 ± 1.34	13.50 ± 15.50	0.94	0.89	0.89
10%	Original	0.95 ± 0.06	2.95 ± 1.83	45.94 ± 25.25	0.87	**0.02**	**< 0.01**
	Gaussian-filtered	0.98 ± 0.02	4.50 ± 3.32	45.75 ± 31.82	0.10	**0.01**	**< 0.01**
	DL-filtered	0.94 ± 0.08	1.13 ± 0.90	9.94 ± 9.28	0.65	0.98	0.99
5%	Original	0.82 ± 0.32	4.12 ± 2.91	65.38 ± 52.79	> 0.99	**0.01**	**< 0.01**
	Gaussian-filtered	0.95 ± 0.08	11.68 ± 7.45	69.38 ± 43.92	0.74	**< 0.01**	**< 0.01**
	DL-filtered	0.96 ± 0.05	1.09 ± 1.23	8.56 ± 10.94	0.71	> 0.99	0.96

P-values denoting significant differences are shown in bold.

The sensitivity, specificity, accuracy, and AUC for bone metastasis detection are shown in Table 6. All images showed 100% sensitivity for bone metastasis detection. Specificity and accuracy decreased with decreasing count percentage for Original and Gaussian-filtered images but were maintained for DL-filtered images. Original and Gaussian-filtered images significantly differed from the AUC of the Reference image at < 50% counts for the Original image and < 25% for the Gaussian-filtered image. However, no significant difference was observed for the DL-filtered image.

Table 6
Comparison with Reference images of sensitivity, specificity, accuracy, and AUC for detection of bone metastases.

Count percentage	Image type	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC	P-value
100%	Reference	100	85.3	90.0	0.99	–
75%	Original	100	76.5	84.0	0.97	0.18
	Gaussian-filtered	100	76.5	84.0	0.98	0.42
	DL-filtered	100	85.3	90.0	0.98	0.53
50%	Original	100	67.6	78.0	0.89	**0.02**
	Gaussian-filtered	100	64.7	76.0	0.97	0.33
	DL-filtered	100	82.4	88.0	0.97	0.45
25%	Original	100	35.3	56.0	0.84	**0.03**
	Gaussian-filtered	100	38.2	58.0	0.92	**0.03**
	DL-filtered	100	82.4	88.0	0.96	0.33
10%	Original	100	5.9	36.0	0.87	**0.02**
	Gaussian-filtered	100	2.9	34.0	0.81	**< 0.01**
	DL-filtered	100	82.4	88.0	0.96	0.35
5%	Original	100	5.9	36.0	0.65	**< 0.01**
	Gaussian-filtered	100	2.9	34.0	0.68	**< 0.01**
	DL-filtered	100	73.5	82.0	0.95	0.27
P-values denoting significant differences are shown in bold.						

4. Discussion

The present study developed a deep-learning model to improve the image quality of low-count whole-body bone scintigraphy. The DL-filtered images output by our model showed better results in PSNR and SSIM than the Original and Gaussian-filtered images. It showed that the diagnostic performance was assured even at a 5% count. We demonstrated the benefit of our developed model, with respect to image quality and diagnostic performance.

In deep learning, a reduction in dimensionality due to encoding deteriorates the output image's resolution. In the present study, incorporating unsharp masking in the model improved resolution and helped depict bone morphology.

The PSNR and SSIM of the DL-filtered image were significantly higher than those of other images. PSNR was improved by signal restoration and noise reduction using deep learning, and SSIM was improved by preserving bone morphology using unsharp masking. Therefore, combining deep learning and unsharp masking led to good performance. On the other hand, the Gaussian-filtered image had lower PSNR and SSIM than the Original image and did not contribute to improved image quality. Ito et al. also showed the same result for the Gaussian-filtered image [16]. Furthermore, Ardenfors et al. [30] noted the risk of signal loss due to smoothing. In the present study, part of the bone signal was lost in Gaussian-filtered images with < 10% counts, and mottled accumulation textures were observed.

DL-filtered images did not differ significantly from Reference images in terms of ANN values, BSI, and Hs at any count percentage, regardless of the presence or absence of bone metastases. In addition, DL-filtered images had higher accuracy in detecting bone metastases than other images and did not differ significantly from Reference images in AUC. Therefore, our model suggests that even low counts may not affect bone metastasis diagnosis.

Original and Gaussian-filtered images significantly affected ANN values, BSI, and Hs at < 25% counts for patients without bone metastases. They significantly affected BSI and Hs at < 10% counts for patients with bone metastases. This was because the noise in the Original images, which increases with lower counts, and the accumulated texture in the Gaussian-filtered images, which becomes more mottled with lower counts, misidentified normal bone as bone metastases. However, the ANN values of the Original and Gaussian-filtered images of patients with bone metastases were not significantly different at any count percentage. This is because high count accumulation at bone metastasis sites improved their recognition. Another reason is that the ANN value is expressed as a value between 0 and 1, with an upper limit. If noise or mottled accumulation of texture due to low counts is misidentified as bone metastases, the ANN value approaches 1; thus, it was unlikely to make a difference in patients with bone metastases whose ANN values in the Reference images were initially close to 1. Therefore, it cannot be argued that the ANN values of patients with bone metastases are unaffected by low-count Original and Gaussian-filtered images.

The sensitivity of the Original and Gaussian-filtered images for detecting bone metastases was 100% at all count percentages; however, the specificity and accuracy decreased considerably with lower count percentages. In other words, false diagnosis of patients without bone metastases was increased with

lower count percentages. The AUC of the Original and Gaussian-filtered images was smaller with decreasing count percentage; in particular, it significantly differed from the Reference image for the Original image with < 50% counts and the Gaussian-filtered image with < 25% counts. Minarik et al. visually evaluated the presence of bone metastases for Gaussian-filtered and DL-filtered images with 50% counts and reported no significant difference from the diagnostic performance of the Reference image [29]. The present study's results supported their findings and further demonstrated the limitations of Original and Gaussian-filtered images by evaluating lower-count percentage images.

Minarik et al. [29] showed the image quality improvement of the deep learning model for 10% counts; however, the current study showed the image quality improvement and the assurance of diagnostic accuracy for the even lower count of 5%. However, the reliability of the output image is vital for image generation by deep learning models, and it is necessary to check whether false information, such as false accumulation or false accumulation loss, is added to the output image. Minarik et al. [29] reported that false accumulation was identified in the 10% count of DL-filtered images. The present study visually identified slight local accumulation differences in the 5% count DL-filtered images. For example, the posterior image of patients without bone metastases showed a slight enhancement of the lower cervical spine (Fig. 2b). The anterior image of patients with bone metastases obscured the accumulation on the right side of the sacroiliac joint (Fig. 3a). The diagnostic performance did not differ significantly from the Reference image; however, it may not be optimal to apply this model to images with only 5% counts.

The interpretation of bone scintigraphy depends on the personal knowledge and experience of the rater. However, CAD software, such as BONENAVI, enables quantitative evaluation and reduces inter-rater variance regardless of the rater's experience [18, 31]. Evaluation criteria for bone scintigraphy may influence the interpretation of results. For example, the visual evaluation by two raters in Minarik et al. [29] used the presence or absence of bone metastases as an evaluation criterion. Although assessing the presence or absence of bone metastases is necessary for the stage diagnosis, it is insufficient for bone metastases follow-up. The results of the present study, supported by the quantitative assessment by BONENAVI analysis, further extend the evidence from Minarik et al.'s visual assessment [29].

In clinical practice, it takes effort to acquire low-count images. Some gamma cameras allow retrospective adjustment of acquisition counts but are limited by the model of the gamma camera in possession [16]. A previous study that directly acquired low-count clinical images with research-specific short-time imaging protocols had a limited number of patients [8]. The approach of mixing different types of images, such as phantom images, into the training images has also been proposed to increase the number of training images [29]. In the present study, it was possible to train the model with many clinical images using the Poisson resampling method, which allows the count percentage to be adjusted retrospectively.

There are several limitations to the present study. First, the dataset used to build the deep learning model was acquired with a single type of gamma camera and dominated by breast cancer patients. The image quality obtained depends on the performance of the gamma camera, and the characteristics of bone

metastases, such as osteogenic and osteolytic types, differ depending on the disease. A multicenter study will be necessary to improve further generalization. Second, BONENAVI analysis was used to evaluate diagnostic performance. BONENAVI analysis can reduce errors associated with the reader; however, the physician's diagnosis is also vital in clinical practice. For clinical use, additional evaluation by the physician should be performed.

5. Conclusion

Our deep learning model has demonstrated good image quality improvement and diagnostic performance for low-count bone scintigraphy, suggesting its applicability to clinical practice. Our model may be used to shorten examination time and reduce the amount of radioactivity administered.

Declarations

References

1. Macedo F, Ladeira K, Pinho F, Saraiva N, Bonito N, Pinto L, et al. Bone Metastases: An Overview. Oncol Rev 2017;11(1):321.

2. Maffioli L, Florimonte L, Pagani L, Butti I, Roca I. Current role of bone scan with phosphonates in the follow-up of breast cancer. Eur J Nucl Med Mol Imaging 2004;31(1):S143-8.

3. Govaert GAM, Glaudemans AWJM. Nuclear medicine imaging of posttraumatic osteomyelitis. Eur J Trauma and Emerg Surg 2016;42(4):397-410.

4. Abdelrazek S, Szumowski P, Rogowski F, Kociura-Sawicka A, Mojsak M, Szorc M. Bone scan in metabolic bone diseases. Review. Nucl Med Rev Cent East Eur 2012;15(2):124-31.

5. Koppula BR, Morton KA, Al-Dulaimi R, Fine GC, Damme NM, Brown RKJ. SPECT/CT in the Evaluation of Suspected Skeletal Pathology. Tomography 2021;7(4):581-605.

6. Saha S, Burke C, Desai A, Vijayanathan S, Gnanasegaran G. SPECT-CT: applications in musculoskeletal radiology. Br J Radiol 2013;86(1031):20120519.

7. Zhang L, He Q, Zhou T, Zhang B, Li W, Peng H, et al. Accurate characterization of (99m)Tc-MDP uptake in extraosseous neoplasm mimicking bone metastasis on whole-body bone scan: contribution of SPECT/CT. BMC Med Imaging 2019;19(1):44.

8. Pan B, Qi N, Meng Q, Wang J, Peng S, Qi C, et al. Ultra high speed SPECT bone imaging enabled by a deep learning enhancement method: a proof of concept. EJNMMI Phys 2022;9(1):43.

9. Pain CD, Egan GF, Chen Z. Deep learning-based image reconstruction and post-processing methods in positron emission tomography for low-dose imaging and resolution enhancement. Eur J Nucl Med Mol Imaging 2022;49(9):3098-118.

10. Cheng Z, Wen J, Huang G, Yan J. Applications of artificial intelligence in nuclear medicine image generation. Quant Imaging Med Surg 2021;11(6):2792-822.

11. Arabi H, Zaidi H. Applications of artificial intelligence and deep learning in molecular imaging and radiotherapy. Eur J Hybrid Imaging 2020;4(1):17.

12. Nensa F, Demircioglu A, Rischpler C. Artificial Intelligence in Nuclear Medicine. J Nucl Med 2019;60(Suppl 2):29s-37s.

13. Shao W, Rowe SP, Du Y. Artificial intelligence in single photon emission computed tomography (SPECT) imaging: a narrative review. Ann Transl Med 2021;9(9):820.

14. Decuyper M, Maebe J, Van Holen R, Vandenberghe S. Artificial intelligence with deep learning in nuclear medicine and radiology. EJNMMI Phys 2021;8(1):81.

15. Zhang D, Pretorius PH, Lin K, Miao W, Li J, King MA, et al. A novel deep-learning-based approach for automatic reorientation of 3D cardiac SPECT images. Eur J Nucl Med Mol Imaging 2021;48(11):3457-68.

16. Ito T, Maeno T, Tsuchikame H, Shishido M, Nishi K, Kojima S, et al. Adapting a low-count acquisition of the bone scintigraphy using deep denoising super-resolution convolutional neural network. Phys Med 2022;100:18-25.

17. Sadik M, Hamadeh I, Nordblom P, Suurkula M, Höglund P, Ohlsson M, et al. Computer-assisted interpretation of planar whole-body bone scans. J Nucl Med 2008;49(12):1958-65.

18. Sadik M, Suurkula M, Höglund P, Järund A, Edenbrandt L. Improved classifications of planar whole-body bone scans using a computer-assisted diagnosis system: a multicenter, multiple-reader, multiple-case study. J Nucl Med 2009;50(3):368-75.

19. Sadik M, Jakobsson D, Olofsson F, Ohlsson M, Suurkula M, Edenbrandt L. A new computer-based decision-support system for the interpretation of bone scans. Nucl Med Commun 2006;27(5):417-23.

20. White D, Lawson RS. A Poisson resampling method for simulating reduced counts in nuclear medicine images. Phys Med Biol 2015;60(9):N167-76.

21. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *arXiv:* 1505.04597.

22. Liu CC, Qi J. Higher SNR PET image prediction using a deep learning model and MRI image. Phys Med Biol 2019;64(11):115004.

23. Chen KT, Gong E, de Carvalho Macruz FB, Xu J, Boumis A, Khalighi M, et al. Ultra-Low-Dose (18)F-Florbetaben Amyloid PET Imaging Using Deep Learning with Multi-Contrast MRI Inputs. Radiology 2019;290(3):649-56.

24. Lu W, Onofrey JA, Lu Y, Shi L, Ma T, Liu Y, et al. An investigation of quantitative accuracy for deep learning based denoising in oncological PET. Phys Med Biol 2019;64(16):165019.

25. Morishita K, Yamagata S, Okabe T, Yokoyama T, Hamatani K. Unsharp masking for image enhancement. United States patent US 4 794 531. 1988.

26. Polesel A, Ramponi G, Mathews VJ. Image enhancement via adaptive unsharp masking. IEEE Trans. Image Process 2000:9(3):505-10.

27. Deng G. A generalized unsharp masking algorithm. IEEE Trans. Image Process. 2011;20(5):1249-61.

28. Ye W, Ma KK. Blurriness-guided unsharp masking. IEEE Trans. Image Process 2018;27(9):4465-77.

29. Minarik D, Enqvist O, Trägårdh E. Denoising of Scintillation Camera Images Using a Deep Convolutional Neural Network: A Monte Carlo Simulation Approach. J Nucl Med 2020;61(2):298-303.

30. Ardenfors O, Svanholm U, Jacobsson H, Sandqvist P, Grybäck P, Jonsson C. Reduced acquisition times in whole body bone scintigraphy using a noise-reducing Pixon®-algorithm-a qualitative evaluation study. EJNMMI Res 2015;5(1):48.

31. Liu S, Feng M, Qiao T, Cai H, Xu K, Yu X, et al. Deep Learning for the Automatic Diagnosis and Analysis of Bone Metastasis on Bone Scintigrams. Cancer Manag Res 2022;14:51-65.
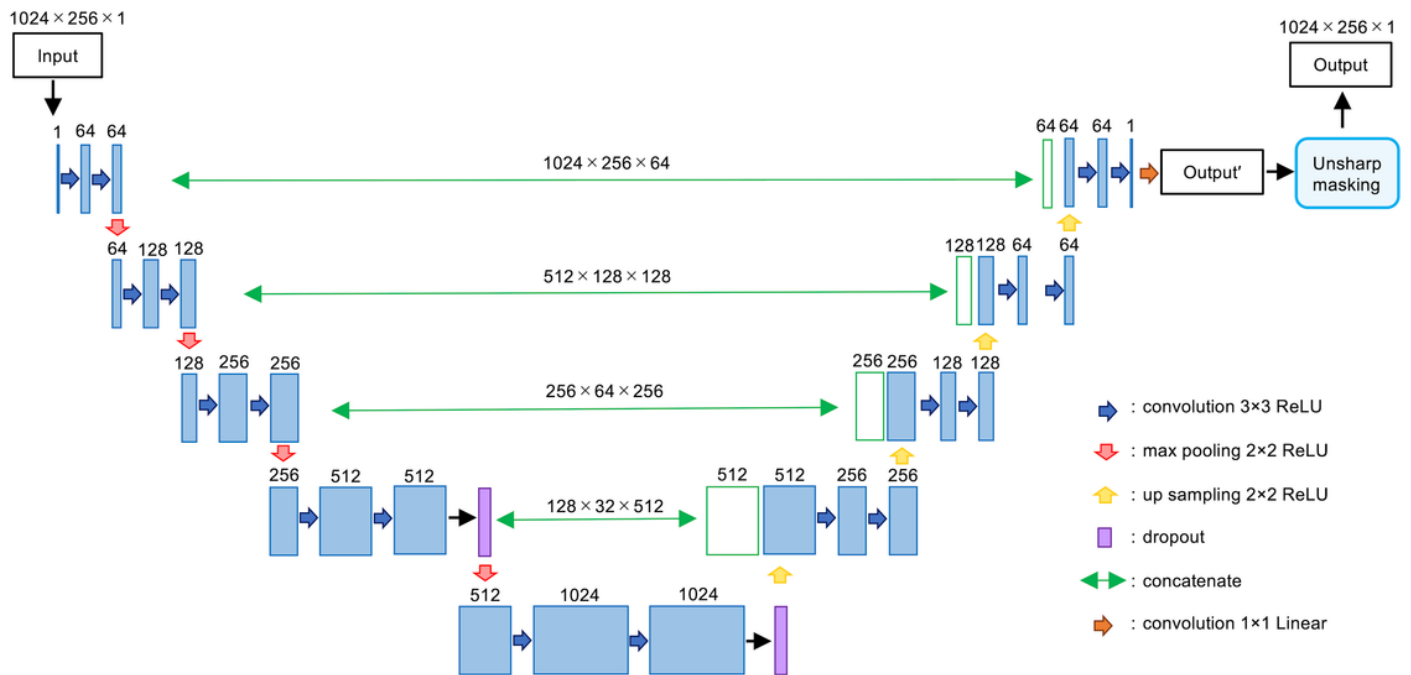
Figures

Figure 1

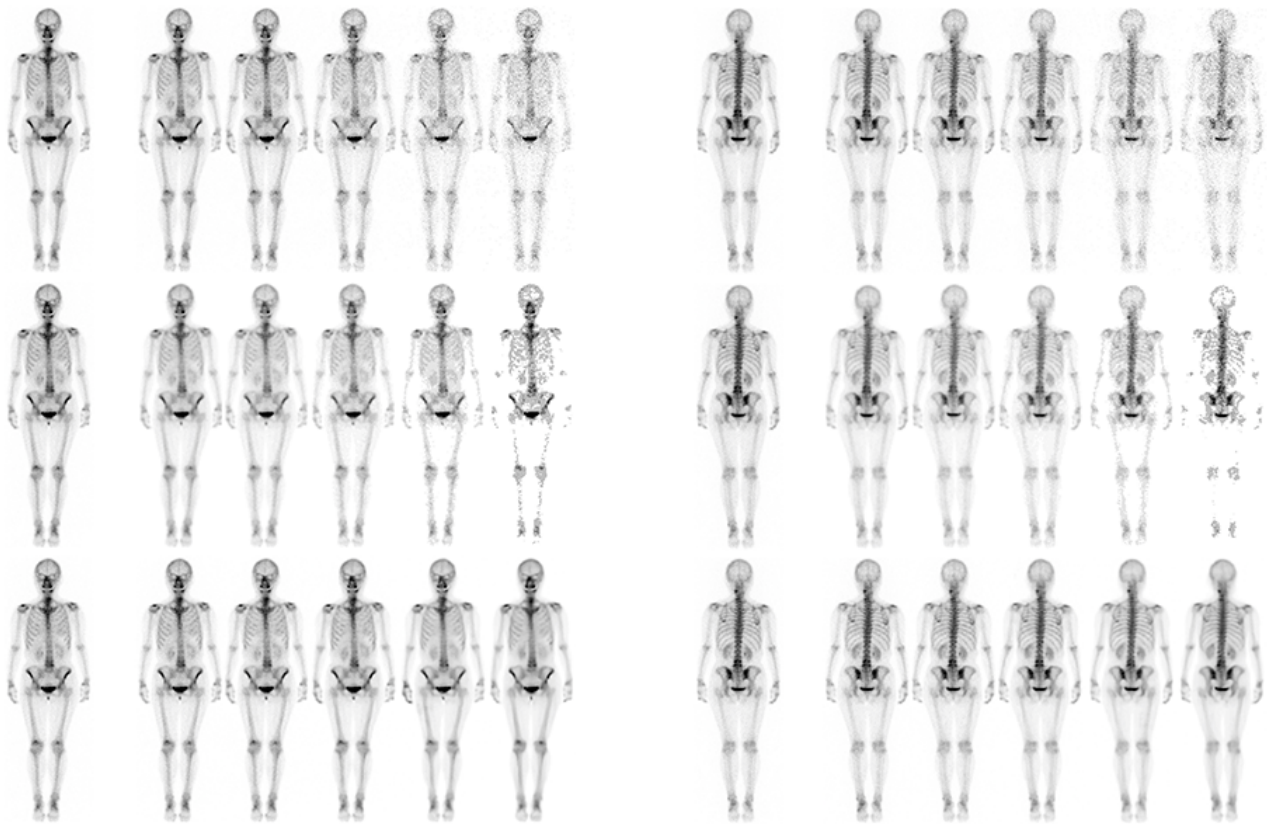Architecture of our deep learning model based on U-Net.

Figure 2

Anterior (a) and posterior (b) bone scintigraphy images of a patient without bone metastases. From top to bottom: Original, Gaussian, and DL-filtered images. From left to right: Reference image (100%), Original image with 75%, 50%, 25%, 10%, and 5% counts of the Reference image.
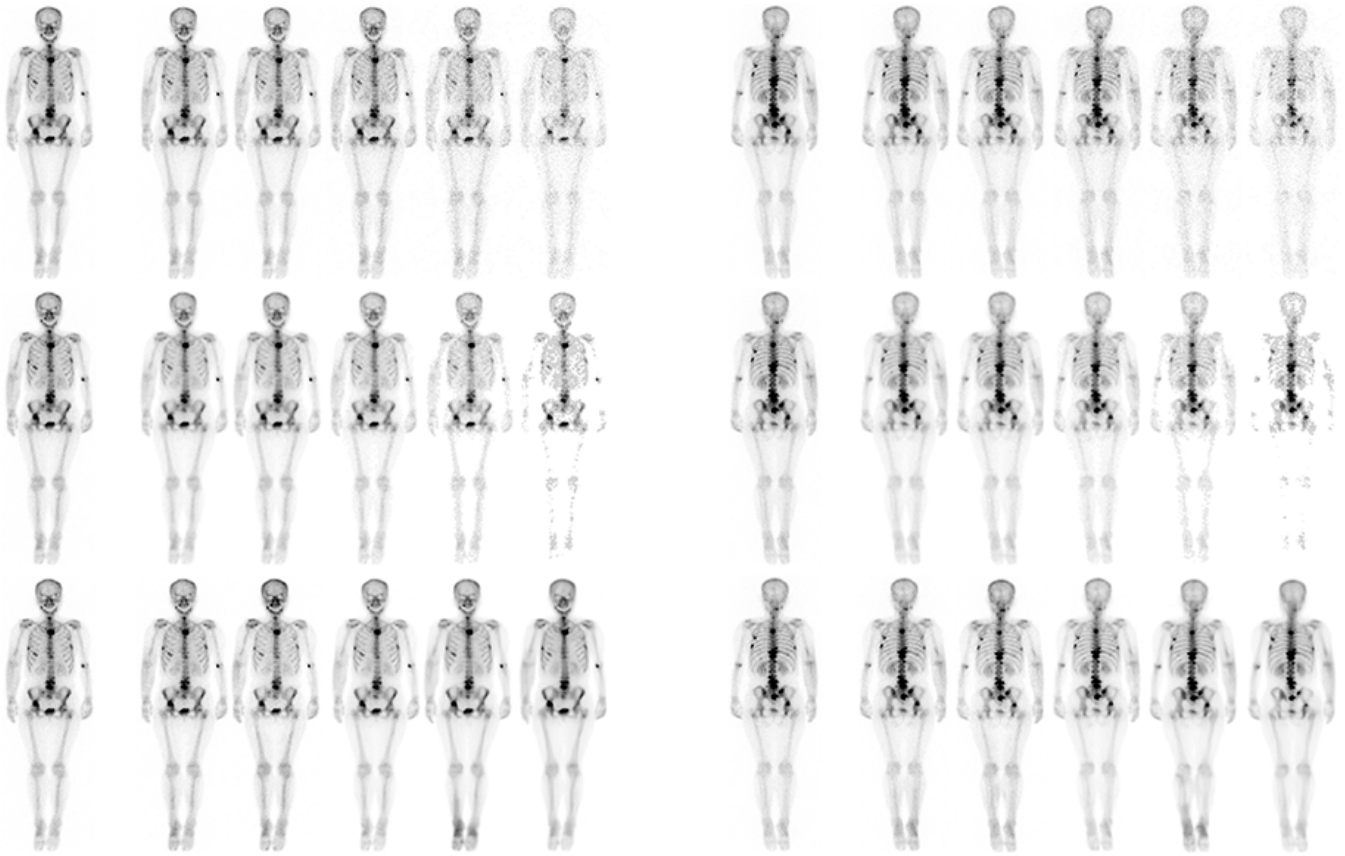
Figure 3

Anterior (a) and posterior (b) bone scintigraphy images of a patient with bone metastases. From top to bottom: Original, Gaussian, and DL-filtered images. From left to right: Reference image (100%), Original image with 75%, 50%, 25%, 10%, and 5% counts of the Reference image.