Enhanced Estimation of Bone Mineral Content from an X-ray Image Using Random Fast Denoising Diffusion Probabilistic Models to Quantify Prediction Uncertainty for Osteoporosis Diagnosis.

Samuel Abena Kpalemon[1] KPALEMON.ABENA_SAMUEL.KC4@NAIST.AC.JP
Yi Gu[1] GU.YI.GU4@NAIST.AC.JP
Naoaki Ono[2,1] NONO@IS.NAIST.JP
Altaf-Ul-Amin[1] AMIN-M@IS.NAIST.JP
Ahmad Kamal Nasution[1] KAMAL.NASUTION@NAIST.AC.JP
Shigehiko Kanaya[1] SKANAYA@IS.NAIST.JP
Keisuke Uemura[4] KEISUKEUEMURA@ORT.MED.OSAKA-U.AC.JP
Mazen Soufi[1] MSOUFI@IS.NAIST.JP
Masaki Takao[6] TAKAO.MASAKI.TI@EHIME-U.AC.JP
Hugues Talbot[7] HUGUES.TALBOT@CENTRALESUPELEC.FR
Seiji Okada[4] SEOKADA@ORT.MED.OSAKA-U.AC.JP
Nobuhiko Sugano[5] N-SUGANO@UMIN.NET
Yoshinobu Sato[3] YOSHI@IS.NAIST.JP
Yoshito Otake[1] OTAKE@IS.NAIST.JP

[1] *Information Science, Graduate School of Science and Technology, Nara Institute of Science and Technology*
[2] *Data Science Center, Nara Institute of Science and Technology*
[3] *Nara Medical University*
[4] *Department of Orthopaedic Surgery, Osaka University Graduate School of Medicine*
[5] *Department of Orthopedic Medical Engineering, Osaka University Graduate School of Medicine*
[6] *Department of Bone and Joint Surgery, Ehime University Graduate School of Medicine*
[7] *CentraleSupélec, Université Paris-Saclay*

Editors: Under Review for MIDL 2025

Abstract

Osteoporosis is the leading cause of bone fractures in the elderly, yet it often goes undiagnosed mainly due to the high cost and limited accessibility of dual-energy X-ray absorptiometry (DXA), the current gold standard for diagnosis. Additionally, deep learning algorithms like Generative Adversarial Networks (GANs) have shown promising results in diagnosing osteoporosis by estimating bone mineral content (BMC). However, previous GAN-based approaches have not addressed uncertainty estimation, which is critical to improving reliability in clinical decision-making. To build on previous work, we propose a novel method that combines random fast denoising diffusion probabilistic model (Random Fast-DDPM) for BMC estimation with a variance-based uncertainty quantification technique. Unlike GANs, which only estimate BMC, our method also captures prediction uncertainty, adding a layer of reliability to the diagnostic process. The proposed method achieved a high Pearson correlation coefficient, r=0.82 in BMC estimation and reported an overall uncertainty score of 0.189 which needs further investigation to assess its reliability.

Keywords: Osteoporosis, Uncertainty, Bone Mineral Content, Diffusion Models.

1. Introduction

Osteoporosis is a fast-growing bone disease and is expected to affect over 263 million people by 2035 (Zhu et al., 2023). Unfortunately, dual-energy X-ray absorptiometry (DXA), is not cost-effective and is often not available to many patients. This creates a big burden for both healthcare providers and patients. At the same time, the field of artificial intelligence has made strong progress in recent years, leading to new solutions such as using Generative Adversarial Networks (GANs) to estimate bone mineral content (BMC) (Gu et al., 2023). While these methods show promise, they lack an essential component for clinical adoption: an estimation of how confident the model is in its predictions. In real-world diagnosis support, especially in borderline or atypical cases, it is crucial not only to generate a prediction but also to convey how much uncertainty surrounds that prediction. Without such information, clinicians may either over-rely on or distrust the model entirely. In addition, diffusion models have recently shown state-of-the-art results in high-quality medical image generation compared to GANs (Müller-Franzes et al., 2023). But, fast denoising diffusion probabilistic models (Fast-DDPM), a newer diffusion approach designed for faster and less computationally intensive image generation than standard DDPM (Jiang et al., 2024), can still produce poor image quality when sampling from random timesteps. Therefore, to combine fast and high-quality image generation which is essential for accurate BMC prediction, with uncertainty quantification, we propose a Random Fast Denoising Diffusion Probabilistic Model (Random Fast-DDPM) with a variance-based uncertainty quantification technique. The key idea is to train the model like a standard DDPM using all 1,000 timesteps, but to sample like Fast-DDPM using uniform or non-uniform random timesteps. This design enables the model to generate high-quality images at any sampling timestep while also allowing straightforward uncertainty estimation.

2. Method

2.1. Dataset and Preprocessing

The dataset comprises 600 patients who underwent hip arthroplasty at Osaka University Hospital (May 2011–Dec 2015). Each patient has 4–5 X-ray images in different poses, paired with a proximal femur (PF-DRR) and corresponding CT-derived BMC. In total, there are 2651 image pairs, with 2118 used for training and 533 for testing. Additionally, a validation set of 120 CT-BMC values is provided in a CSV file. For preprocessing, all images were resized from 256×128 to 128×64 and normalized to the [0, 1] range.

2.2. Model Architecture

The proposed model is based on a U-Net architecture composed of encoder, bottleneck, and decoder modules. The input is a concatenation of the PF image and its corresponding X-ray (2 channels), with batch size B=8, and is projected to shape B×128×128×64 by a convolutional layer. The encoder has 6 downsampling blocks, each comprising two ResNet blocks with normalization, attention, activation, timestep embedding, and downsampling layers. This reduces spatial resolution to 4×2 and increases channel depth to 512. The bottleneck contains two mid-blocks, each with two ResNet layers, attention, and embedding layers, maintaining shape B×512×4×2 to enhance global context. The decoder reverses

the encoder and uses skip connections to progressively upsample back to $B\times128\times128\times64$, followed by a final convolutional layer producing the output shape $B\times1\times128\times64$.

2.3. Training and Sampling Protocol

To run Algorithm 1, we trained the model on a NVIDIA A6000 GPU using a linear scheduler to gradually add noise. We used MSE as the loss to compare predicted and true noise, and optimized with Adam at a learning rate of 0.00002.

Table 1: Training and Sampling Procedures for Random Fast DDPM

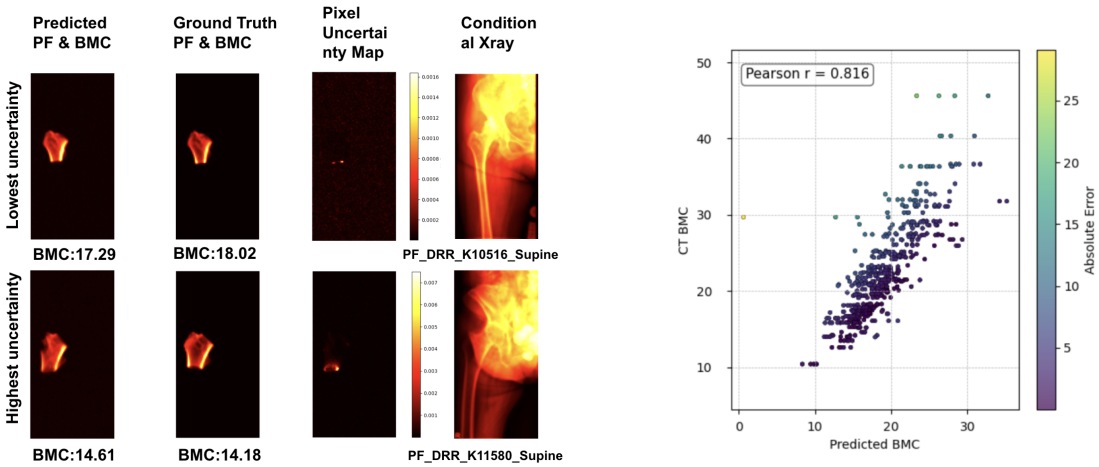
Training Procedure (Algorithm 1)	**Sampling Procedure (Algorithm 2)**
1. Initialize	**1. Initialize**
Shuffle timesteps: $S \leftarrow$ Shuffle$(\{0,\dots,T-1\})$, set $i \leftarrow 0$	Sample conditional image $c \sim p_c(c)$ and noise $\epsilon \sim \mathcal{N}(0,\mathbb{I})$
2. Sample and Prepare	**2. Add Noise**
Sample training pair $(x_0, c) \sim p_{\text{joint}}(x_0, c)$	Compute noisy input: $x(T) = \alpha(T)c + \sigma(T)\epsilon$
If $i \geq T$, reshuffle S, reset $i \leftarrow 0$	Select timesteps $\{t_1,\dots,t_N\} \subset \{T-1,\dots,0\}$ (e.g., 10 steps)
Choose timestep: $t \leftarrow S[i]$, then $i \leftarrow i+1$	
3. Noise Injection	**3. Iterative Denoising**
Sample noise $\epsilon \sim \mathcal{N}(0,\mathbb{I})$	For each t in $\{t_1,\dots,t_N\}$:
Compute: $\tilde{x} = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon$	Form $\hat{x}(t) = [x(t) \parallel c]$
Concatenate: $\hat{x} = [\tilde{x} \parallel c]$	Predict noise: $\epsilon_\theta(\hat{x}(t), t)$
	Update: See Eq. 1
4. Optimization	**4. Output**
Update model: $\nabla_\theta \|\epsilon - \epsilon_\theta(\hat{x}, t)\|^2$	Return final result: $x(0)$

$$x(t-\Delta t) = \frac{\alpha(t-\Delta t)}{\alpha(t)}x(t) + \left[\sigma(t-\Delta t) - \frac{\alpha(t-\Delta t)}{\alpha(t)}\sigma(t)\right]\epsilon_\theta(\hat{x}(t), t) \tag{1}$$

Algorithm 1 handles training by adding noise to PF images and learn to denoise using shuffled timesteps. Algorithm 2 performs sampling and estimate BMC to quantify uncertainty.

3. Results

In this first experiment, our proposed Random Fast-DDPM used the same uniform timesteps as Fast-DDPM (Jiang et al., 2024), specifically `[0, 199, 299, 399, 499, 599, 699, 799, 899, 999]`, and achieved a Pearson correlation of **r=0.82** for BMC estimation (Figure 1b), demonstrating strong reconstruction capability from any timestep. Figure 1a shows the conditional X-ray and predicted PFs with the **lowest** and **highest uncertainty** maps, corresponding to BMC values of **17.28** and **14.61**, respectively. These maps, based on pixel-wise variance, highlight areas of low model confidence. Overall, BMC prediction uncertainty was low, with a mean of **0.189**, ranging from **0.016** to **1.176**. While the model is



(*a*) PFs, BMCs, and uncertainty maps. (*b*) (BMC p.c.c : $r = 0.82$).

Figure 1: Qualitative and quantitative evaluation of BMC prediction.

generally confident, further analysis is needed to evaluate how uncertainty relates to actual prediction error, in order to fully validate our uncertainty estimation method.

References

Yi Gu, Yoshito Otake, Keisuke Uemura, Mazen Soufi, Masaki Takao, Hugues Talbot, Seiji Okada, Nobuhiko Sugano, and Yoshinobu Sato. Bone mineral density estimation from a plain x-ray image by learning decomposition into projections of bone-segmented computed tomography. *Medical Image Analysis*, 90:102970, 2023.

Hongxu Jiang, Muhammad Imran, Linhai Ma, Teng Zhang, Yuyin Zhou, Muxuan Liang, Kuang Gong, and Wei Shao. Fast-ddpm: Fast denoising diffusion probabilistic models for medical image-to-image generation. *arXiv preprint arXiv:2405.14802*, 2024.

Gustav Müller-Franzes, Jan Moritz Niehues, Firas Khader, Soroosh Tayebi Arasteh, Christoph Haarburger, Christiane Kuhl, Tianci Wang, Tianyu Han, Teresa Nolte, Sven Nebelung, et al. A multimodal comparison of latent denoising diffusion probabilistic models and generative adversarial networks for medical image synthesis. *Scientific Reports*, 13(1):12098, 2023.

Z Zhu, P Yu, Y Wu, Z Tan, J Ling, J Ma, Jing Zhang, W Zhu, and Xiao Liu. Sex specific global burden of osteoporosis in 204 countries and territories, from 1990 to 2030: an age-period-cohort modeling study. *The Journal of nutrition, health and aging*, 27(9):767–774, 2023.