SIIM23 ANNUAL MEETING + InformaticsTECH Expo
JUNE 14-16 | AUSTIN, TX
Learning. Innovating. Belonging.

High-Fidelity Conditional Pelvis Radiograph Generation with Denoising Diffusion Probabilistic Models

Bardia Khosravi, MD, MPH, MHPE, Postdoctoral Research Fellow, Mayo Clinic
Pouria Rouzrokh, MD, MPH, MHPE; John Mickley, MD; Shahriar Faghani, MD, MSc; Bradley Erickson, MD, PhD; Michael Taunton, MD; Cody Wyles, MD

Introduction

Recent advances in deep learning have pushed the boundaries of image generation. Generative adversarial networks (GANs) are the most famous designs for creating realistically looking synthetic images. However, GANs do not capture the full data distribution and suffer from mode-collapse. Recently, denoising diffusion probabilistic models (DDPMs) have been introduced as an alternative to GANs to create diverse images with even better quality. The famous text-to-image model Stable Diffusion uses a DDPM-based approach to create high-fidelity images based on text prompts.

Hypothesis

We hypothesize that DDPMs can be used to capture nuances of data distribution and can be conditioned to create various images based on specified input text.

Methods

A set of 480,402 pelvis radiographs captured between 1998 and 2018 were gathered from an institutional registry. Images were labeled based on their view (AP/lateral/oblique), joint side present (right/left/both), presence of prosthesis on each side (yes/no), and patient's demographic information (age, sex, and BMI). A DDPM with a cosine noise schedule was used with a total of T=1000 steps. To train this model, every image in a batch is converted to a "noisy" version by adding a specific amount of noise (noise intensity comes from t; 1< t< T) based on Markov chain theory (called forward diffusion). The DDPM tries to denoise the image to a less noisy version (with a noise level of t-1) through a process called reverse diffusion (Figure 1). Classifier-free guidance was used to condition the generated images based on the extracted labels. Two orthopedic surgeons evaluated 100 synthetics for discrepancies between the given labels and image. Additionally, they were provided with another set of 200 images (100 real and 100 synthetics) to evaluate the realism of the generated images. An interactive demo of the model is available at https://demo.osail.ai/Unsupervised_DDPM_Hip.

Results

The image generation model was trained after seeing 40,000,000 images which took around 420 A100 GPU hours. The inference time for creating one thousand 1024*1024 images was 26 minutes. The model reached an FID-5K of 10.4. Reviewers did not find any discrepancy between the synthesized images and the designated labels. Additionally, they had an accuracy of 62% in discriminating real vs. synthetic.

Conclusion

DDPMs can be used for generating high-fidelity images with many conditional restraints, as they can efficiently capture the underlying data distribution. These can be further utilized to create images from minority groups with lower representation in a dataset to have fair training for downstream tasks.

Figure(s)

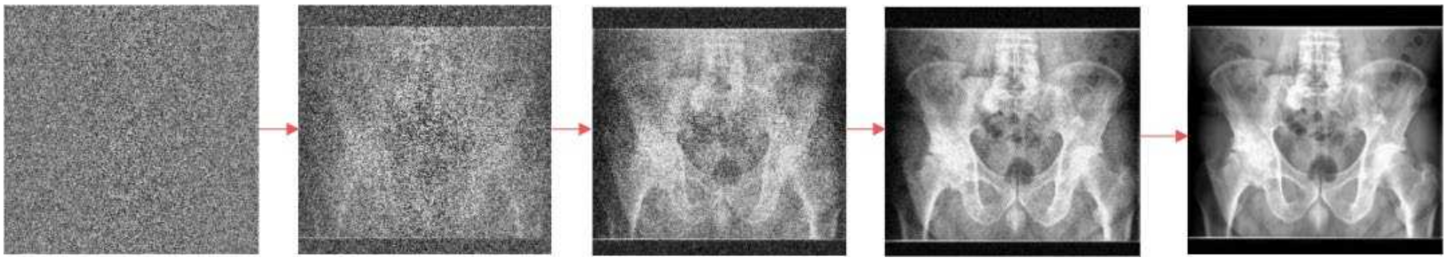
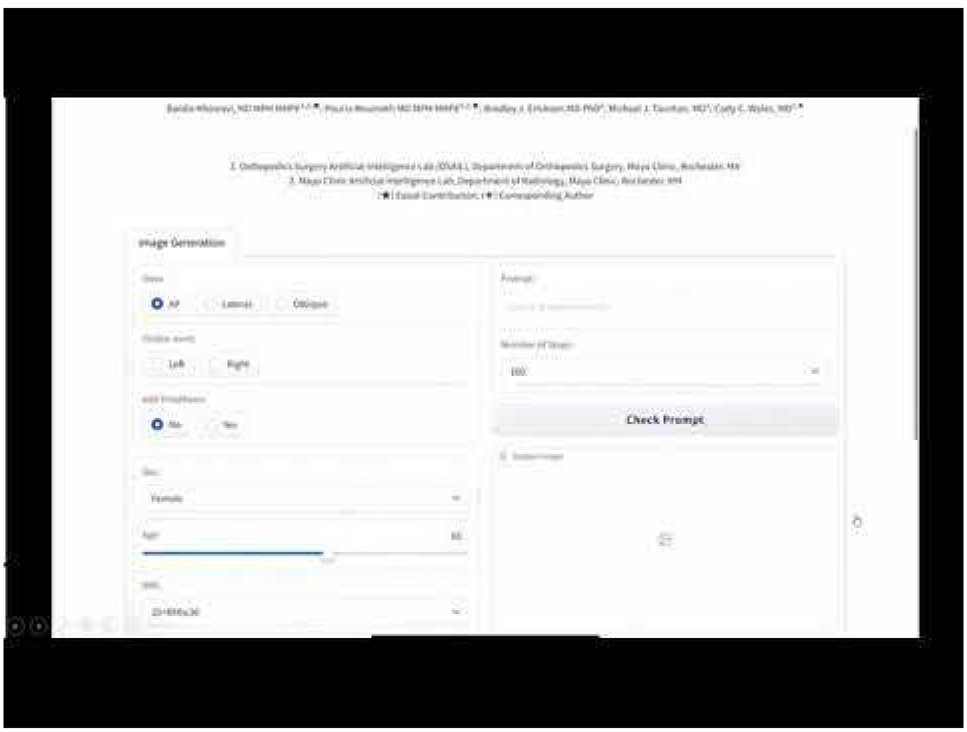


Figure 1. Reverse Diffusion Process Done By a Denoising Diffusion Probabilistic Model (DDPM)



Demo 1. Video demo of the model and its different conditions.

Keywords

Artificial Intelligence/Machine Learning; Emerging Technologies