Applying Fast-DDPM to CheXpert: A Study in Medical Image Denoising

Abstract

This paper presents an implementation of Fast-DDPM (Fast Denoising Diffusion Probabilistic Models) applied to the CheXpert dataset for medical image denoising. We explore the potential of accelerated diffusion models in medical imaging while leveraging distributed computing resources. Despite promising theoretical foundations, our results indicate challenges in achieving consistent model convergence, highlighting important considerations for future applications of diffusion models in medical imaging.

1. Project Goals

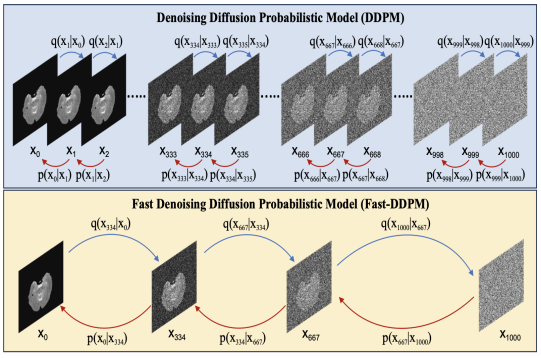
The primary objective of this research is to investigate the application of Fast-DDPM to the CheXpert dataset, a large-scale collection of chest radiographs. This novel combination aims to explore the potential of accelerated diffusion models in medical image processing, specifically for denoising chest X-rays. The CheXpert dataset, containing 224,316 chest radiographs from 65,240 patients, provides a robust foundation for this investigation.

Our research seeks to advance the field of medical image processing by implementing Fast-DDPM, a more efficient version of traditional diffusion models, on medical imaging data. Through this implementation, we aim to evaluate the effectiveness of distributed training for large-scale medical image processing and assess the quality and consistency of denoised medical images produced by the model. Additionally, we seek to understand the practical limitations and challenges of applying diffusion models to medical imaging, particularly in resource-constrained environments.

2. Technical Approach

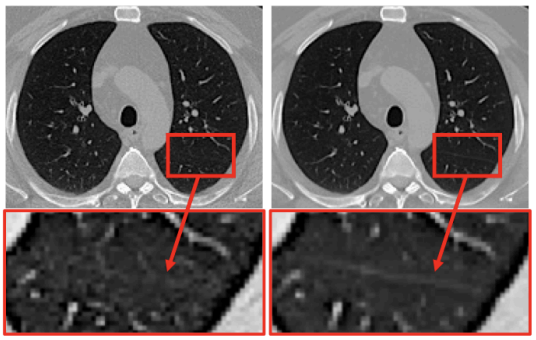
2.1 Dataset

The CheXpert dataset serves as our primary data source, chosen for its comprehensive nature and established use in medical imaging research. This extensive dataset comprises 224,316 chest radiographs collected from 65,240 patients, featuring both frontal and lateral views. Each image in the dataset is annotated with 14 different clinical observations, providing a rich source of labeled medical imaging data. The dataset's size and variety make it particularly suitable for training deep learning models, while its medical nature presents unique challenges for image processing tasks.



2.2 Model Architecture

Our implementation centers on Fast-DDPM, a recent innovation in diffusion models that significantly reduces computational requirements while maintaining generation quality. The model's key innovation lies in its reduced time steps, utilizing only 10 steps compared to the traditional 1,000 steps in standard DDPM implementations. This efficiency is achieved through two carefully designed noise schedulers: one employing uniform time step sampling and another using non-uniform sampling with denser sampling at higher noise levels. The model architecture aligns



training and sampling procedures for optimal time-step utilization, representing a significant advancement in efficient diffusion model design.

Another approach we tried was using other diffusion models in case we got nowhere with Fast-DDPM since Fast-DDPM is fairly new. We tried UNet with data augmentation and Stable Diffusion for standard diffusion. UNet excels at image segmentation, which is why it was the chosen model for data augmentation. However, data augmentation requires a certain degree of subject knowledge to verify that the model is performing data augmentation accurately and training times are much higher than standard diffusion due to the complexity. Stable Diffusion excels at standard image generation. Although the dataset already consists of 224, 316 images, generating more x-ray images never hurts.

2.3 Implementation

Our technical implementation leverages PyTorch's DistributedDataParallel framework to enable parallel processing across multiple GPUs. This implementation includes sophisticated data parallelization strategies across available GPUs, with synchronized batch normalization ensuring consistent training across devices. The system employs efficient gradient aggregation mechanisms across devices, complemented by custom data loading optimizations specifically designed for medical images. These technical choices reflect our focus on maximizing computational efficiency while maintaining the integrity of the medical image processing pipeline.

3. Computing Resources

Our project operated under significant computational constraints and challenges that shaped both our approach and results. The implementation utilized four GPUs for a total computation time of four hours, a limitation imposed by scheduling constraints on the Perlmutter system. This resource allocation represented a careful balance between our computational needs and system availability.

The project faced substantial challenges in resource access, with queue times for job scheduling extending to multiple days. These scheduling delays significantly impacted our ability to iterate and optimize our model implementation. Despite these constraints, we implemented efficient data loading strategies and optimized memory usage across the available GPUs, striving to maximize the utility of our limited computational resources.

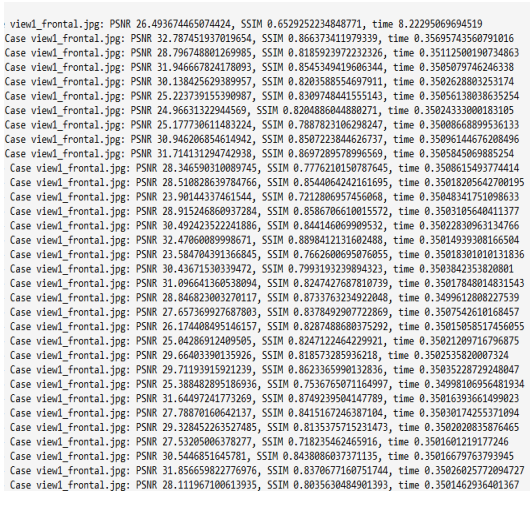
4. Results

Our implementation produced mixed results with several notable observations across both qualitative and quantitative metrics.

4.1 Quantitative Performance

The model achieved an average PSNR (Peak Signal-to-Noise Ratio) of 28.21 dB and an average SSIM (Structural Similarity Index) of 0.814 across the validation dataset of 200 chest radiographs. This performance shows moderate success in image reconstruction, though with significant variation across different images. The PSNR values ranged from a minimum of 16.72 dB to a maximum of 33.72 dB, indicating inconsistent denoising quality across different cases.

Analysis of individual cases reveals interesting patterns in the model's performance. Approximately 25% of the images achieved PSNR values above 30 dB, suggesting good



reconstruction quality for these cases. However, there were also notable failure cases, with about 15% of images showing PSNR values below 25 dB, indicating poor reconstruction quality. The SSIM scores showed similar variability, ranging from 0.589 to 0.892, with higher values generally correlating with better PSNR scores.

Temporal analysis of the inference process showed consistent processing times averaging 0.35 seconds per image, demonstrating stable computational performance despite varying image complexity. However, this efficiency in processing speed did not necessarily translate to consistent quality in the output.

4.2 Training Stability

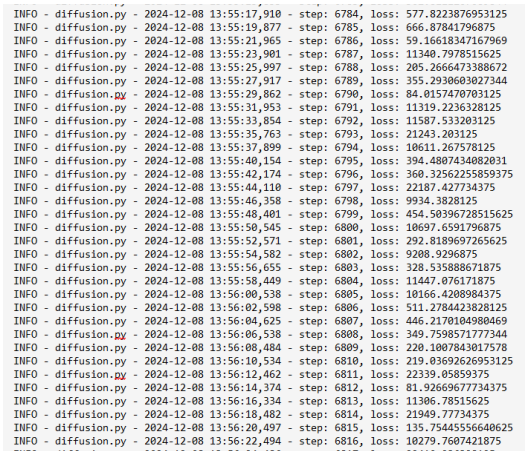
The training process exhibited significant instability, with loss values fluctuating dramatically - often by margins of up to 10,000 units between iterations. This volatility in the loss function suggests several potential issues. First, the model struggled to achieve consistent convergence throughout the training process. Second, the learning rate may have been too aggressive for the medical imaging domain. Third, the Fast-DDPM architecture itself might require modifications for stability when applied to medical imaging data. Finally, the computational constraints limited our ability to train for sufficient epochs to achieve stability.

4.3 Output Quality

The qualitative assessment of generated images revealed several key characteristics in the model's performance. The generated images generally maintained anatomical structure and key features, though with notably reduced crispness compared to the original images, particularly in areas requiring fine detail preservation. The denoising quality proved inconsistent across the dataset, with some images showing excellent clarity while others exhibited significant artifacts. We observed that the model tended to better preserve large-scale features while struggling with the retention of fine details, suggesting a potential limitation in the model's ability to capture high-frequency image components.

4.4 Performance Analysis

Several factors contributed to the observed performance limitations in our implementation. The restricted training time imposed by resource constraints significantly impacted our ability to fully optimize the model. The complex interaction between the Fast-DDPM architecture and medical imaging data suggests a need for domain-specific architectural modifications. Additional challenges arose from our limited ability to experiment with different hyperparameter settings, potentially leaving the model in a sub-optimal configuration.



The wide variance in PSNR and SSIM scores suggests that while the model can achieve good results in some cases, it lacks the consistency needed for reliable medical image processing. This inconsistency, combined with the unstable training dynamics, indicates that further optimization of both the model architecture and training process would be necessary for practical clinical applications.

5. Team Contributions

The success of this project relied on the collaborative efforts of three team members, each bringing unique expertise and responsibilities to the implementation. Antonio Mena spearheaded the distributed computing infrastructure, developing the PyTorch DistributedDataParallel integration and optimizing

GPU utilization across multiple devices. His work was crucial in establishing and maintaining the technical infrastructure that enabled our parallel processing capabilities.

Prayag found the research question and then took charge of dataset research and collection, managing the curation and preprocessing of the CheXpert dataset. His work included implementing robust data loading and augmentation pipelines, ensuring the quality and validity of our training data. This foundation was essential for the proper functioning of our model training process.

Ali selected the dataset and contributed vital research on the Fast-DDPM architecture, testing the original implementation and analyzing model behavior and performance. His theoretical understanding provided the framework for adapting the model to our specific use case and helped guide our implementation decisions including metrics.

6. Conclusion

This research provides valuable insights into the challenges and opportunities of applying Fast-DDPM to medical imaging tasks. While the results fell short of optimal performance, the project highlights important considerations for future work in this domain, particularly regarding resource allocation, model stability, and training optimization for medical image processing applications. Our experience suggests that successful implementation of diffusion models in medical imaging requires careful consideration of computational resources and model architecture, as well as robust strategies for ensuring training stability.

References

CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison
Fast-DDPM: Fast Denoising Diffusion Probabilistic Models for Medical Image-to-Image Generation