Generative AI-Driven Synthesis for Powder Bed Fusion Melt Pool Images

Uchit Shriyan
School of Manufacturing Systems and Networks

Dr. Hyunwoong Ko
School of Manufacturing Systems and Networks

Spring 2024

1 Abstract

This report presents the development and implementation of a Denoising Diffusion Probabilistic Model (DDPM) tailored for generating synthetic images of melt pools in Powder Bed Fusion (PBF) additive manufacturing processes. The study focuses on leveraging the inherent capabilities of DDPMs to simulate and predict the complex thermal and fluid dynamics occurring within melt pools under varied processing conditions. By integrating melt pool images, the model aims to enhance the predictive accuracy and operational efficiency of PBF systems. This research not only contributes to the theoretical understanding of generative models in high-resolution image synthesis but also explores practical applications in optimizing additive manufacturing processes, potentially leading to significant improvements in material properties and part reliability. You can find our code implementation at `https://github.com/UchitShriyan/DDPM_MeltPool`

2 Introduction

2.1 Background

Powder Bed Fusion (PBF) is a prominent additive manufacturing (AM) technique characterized by its ability to produce complex geometries and superior mechanical properties. The quality of the manufactured parts heavily depends on the stability and characteristics of the melt pool formed during the laser or electron beam interaction with the powder material. Traditional approaches to monitoring and controlling melt pool characteristics often rely on trial-and-error adjustments of process parameters, which can be inefficient and yield inconsistent results.

2.2 Problem Statement

Despite advancements in PBF technology, there remains a significant challenge in predicting the outcomes of varying process parameters on the melt pool and thus on the final product quality. Current methods often require extensive experimentation and real-time adjustments, which are not always feasible or economically viable.

2.3 Objectives

The primary objective of this project is to develop a generative model capable of simulating melt pool dynamics that reflect realistic physical phenomena. The model aims to predict melt pool images based on existing melt pool image dataset, thereby providing a tool for better understanding of the PBF process.

2.4 Scope

The scope of this project includes:

- Developing a DDPM to generate high-fidelity images of melt pools.
- Integrating real-world data of melt pool images to train the model.
- Assessing the model's effectiveness in generating melt pool images.
- Exploring the potential of the model to serve as a decision-support tool in real-time manufacturing settings.

2.5 Significance

This project holds significant potential to impact the field of additive manufacturing by providing a predictive modeling tool that enhances the understanding and control of melt pool dynamics. By improving the predictability of the manufacturing process, the model can help reduce waste, increase efficiency, and improve the mechanical properties of the manufactured parts. Furthermore, this research contributes to the broader field of applied machine learning by demonstrating the utility of diffusion models in complex, real-world applications.

3 Literature Review

The evolution of generative models has significantly transformed the landscape of image synthesis. Among the spectrum of these models, diffusion models stand out due to their ability to generate high-quality images through a process that simulates thermodynamic diffusion. This approach incrementally adds and then methodically reverses noise, creating finely detailed images. Ho et al. [1] elaborate on the underlying principles of diffusion models, highlighting their derivation from nonequilibrium thermodynamics and their application in producing state-of-the-art image synthesis results.

Diffusion models differ from other prevalent generative models like Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) in several key aspects. While GANs often suffer from training instabilities and mode collapse, diffusion models offer a more structured and stable training process. This structured approach enables the progressive refinement of images, allowing for incremental enhancements in detail and realism [1]. This characteristic makes them particularly suitable for applications requiring high fidelity in image generation, such as in digital art and medical imaging.

The versatility of diffusion models extends to complex applications such as enhancing robotic vision systems. Kapelyukh et al. [2] discuss the integration of web-scale diffusion models into robotics, enabling robots to perform tasks that rely on nuanced visual understanding, such as object rearrangement based on visual inputs. This integration demonstrates the practical applicability of diffusion models beyond mere image generation, venturing into areas that require interaction with physical environments.

However, the computational demands of training diffusion models are significant, often requiring extensive resources which may limit their deployment in resource-constrained settings [1]. Furthermore, maintaining diversity in the outputs of diffusion models without encountering the common pitfalls of generative models, like mode collapse, is an ongoing challenge.

The future of diffusion models appears promising, especially with the potential for integration with other advanced AI technologies. For instance, combining diffusion models with reinforcement learning could lead to the development of more dynamic and responsive AI systems, capable of adapting to changing environments [2]. Such advancements could revolutionize fields ranging from autonomous vehicles to adaptive content creation in digital media.

4 Methodology

4.1 Architecture

The code utilizes a Denoising Diffusion Probabilistic Model (DDPM), which is a class of generative models operating through a Markov chain of diffusion steps defined by Gaussian transitions. The model architecture systematically introduces Gaussian noise into an image and then employs a reverse generative process to reconstruct the original data from the noise.

- **Diffusion Process**: The diffusion or the forwards process is described by a sequence where each step gradually adds noise to the image. The noise

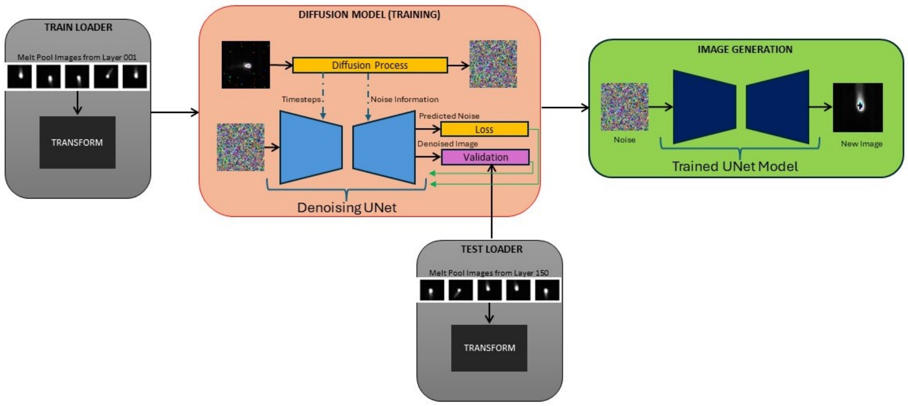

Figure 1: Model Architecture

addition at each timestep t is given by:

$$q(x_t \mid x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}\, x_{t-1}, \beta_t \mathbf{I})$$

where β_t represents the variance of the noise added at each step. The variance parameters β_t are predetermined and progressively increase noise in the data distribution, transitioning it towards a Gaussian distribution over the course of many timesteps.

- **Reverse Process**: The reverse process involves a parameterized model, typically a neural network, which attempts to reconstruct the original image from its noisy counterpart. The model uses the noisy image at timestep t to predict the noise and reverse its effects:

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \sigma_t^2 I)$$

Here, μ_θ is a function modeled by the neural network that predicts the mean of the distribution from which the original image can be reconstructed, effectively learning to denoise the image.

- **U-Net Architecture**: The U-Net model employed in this system is crucial for its ability to handle the complexities of image data, with its architecture featuring a series of downsampling (encoder) and upsampling (decoder) layers connected by skip connections. These connections help the network to retain and utilize contextual information from the input image at various resolution scales, which is critical for accurate reconstruction of details in the output.

- **Sinusoidal Position Embeddings and Block Modules**: These components are integrated into the U-Net architecture to introduce temporal dynamics and label-specific features into the diffusion process, enhancing the model's ability to generate contextually accurate images.

4.2 Computational Framework

To train and optimize this complex architecture, several computational strategies are employed:

- **Loss Function**: The loss function used is the Mean Squared Error (MSE) between the predicted noise and the actual noise in the images. This choice is motivated by the need to precisely match the statistical properties of the noise model, crucial for the reverse diffusion process to accurately reconstruct the original images.

- **Optimization Strategy**: The Adam optimizer is selected for its advantages in handling non-stationary objectives and its efficiency in large-scale data contexts. The learning rate is set to 0.001, providing a balance between rapid convergence and stability of updates.

5 Implementation

5.1 Development Environment and Tools

The code is implemented using Python and use PyTorch for its dynamic graph capabilities and strong GPU acceleration, which are essential for training deep learning models. Torchvision aids in image manipulation, crucial for preprocessing tasks. NumPy is used for numerical operations on arrays, and Matplotlib for visualizing results, especially during the debugging and testing phases.

5.2 Data Handling and Preprocessing

Data handling is managed through custom dataset classes, CustomDataset and MPMDataset, extending PyTorch's torch.utils.data.Dataset. These classes are responsible for loading images from disk, converting them to grayscale to reduce complexity, resizing them to a uniform dimension (64x64 pixels), and normalizing their pixel values to the range [-1, 1] for optimal neural network performance. This preprocessing pipeline ensures that the input data is standardized and ready for efficient processing by the neural network.

5.3 Model and Training Procedure

The architecture of the model is based on a U-Net design, which is pivotal for tasks that involve image-to-image prediction like in this project, where the model predicts noise levels to perform image denoising. The U-Net architecture features:

- **Encoder and Decoder Paths**:The encoder uses convolutional layers to capture the hierarchical features of the images, while the decoder uses transposed convolutions to reconstruct the image from the encoded representation.

- **Skip Connections**: These are crucial as they allow the transfer of fine-grained details from the encoder to the decoder, helping in reconstructing high-quality images.

Each layer includes batch normalization and ReLU activation functions to stabilize and speed up the training process. The model employs an Adam optimizer with a learning rate of 0.001, optimizing the training with efficient computation of adaptive learning rates for different parameters. The training process involves:

- **Noise Simulation and Model Prediction**: In each epoch, images undergo a forward diffusion process to simulate noise addition, followed by a reverse process where the model predicts the noise and reconstructs the input image.

- **Batch Processing**: Utilizing batches of 256 images, leveraging GPU acceleration

- **Epoch and Validation Cycles**: The model is iteratively trained and validated over 150 epochs to ensure it generalizes well over unseen data while avoiding overfitting.

5.4 GPU Utilization and Memory Management

Training is performed on GPU(s) to exploit parallel processing capabilities, significantly reducing the training time. Memory management is handled by periodic calls `torch.cuda.empty_cache()`, which clears unused memory to prevent out-of-memory errors on the GPU, ensuring smooth training sessions.

6 Results

The model was first tested on a single image from the Melt Pool image dataset.

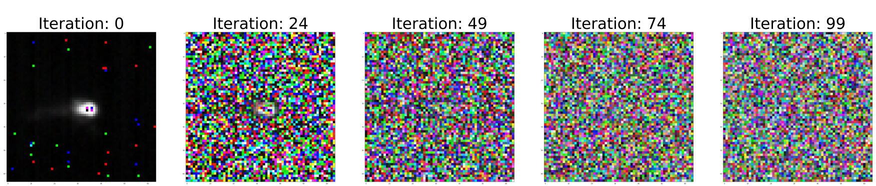

Figure 2: Progression of images generated by the Denoising Diffusion Probabilistic Model at various iterations. Starting with the original image, subsequent images show the addition of noise as the model progresses through the iterations. The final iteration demonstrates the model's capacity to maintain the structural integrity of the melt pool despite the incremental introduction of noise, indicating effective learning of the underlying data distribution.

As seen in Figure 2, the image sequence demonstrates the model's diffusion process over a series of iterations. Initially, the true melt pool image is visible with minimal noise. As the iterations progress, the model incrementally adds noise, simulating the forward diffusion process. By the final iteration, the model exhibits a noisy version of the image that retains the melt pool's core structure, which the reverse diffusion process will aim to reconstruct. This iterative noise addition is crucial for the DDPM to learn the distribution of the data, and the ability of the model to preserve the melt pool's structural integrity throughout this process is indicative of its potential for generating accurate simulations of melt pool dynamics.

A part of the training is shown here to show the loss between the predicted and actual noises to train the UNet:

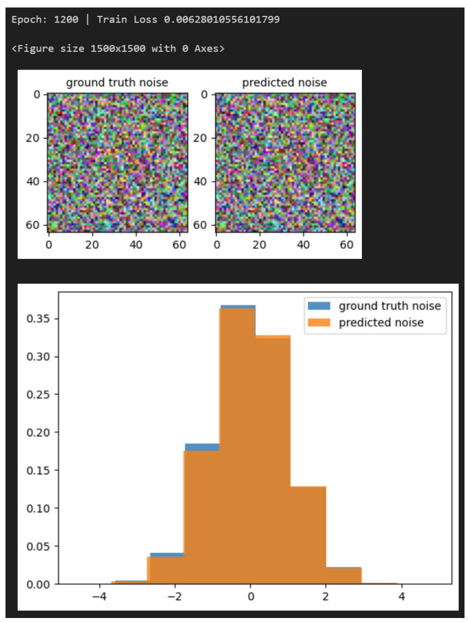

Figure 3: Comparison of ground truth and predicted noise distributions during model training at epoch 1200. The left panel shows the ground truth noise added to an example melt pool image, and the right panel shows the noise predicted by the DDPM. The histogram illustrates the distributions of both noise types, with the ground truth noise in blue and predicted noise in orange, indicating the model's capability to closely approximate the true noise distribution.

As depicted in Figure 3, the outputs reveal the model's proficiency in noise prediction as it trains. At epoch 1200, the DDPM has learned to generate noise that visually and statistically resembles the ground truth noise, as evidenced by the overlapping histograms. This convergence of distributions is indicative of the model's increasing accuracy in noise characterization, a crucial component for successful denoising and image generation in the reverse diffusion process.

For the reverse or denoising process, we wanted to check if the model is able to generate a single image from random noise.

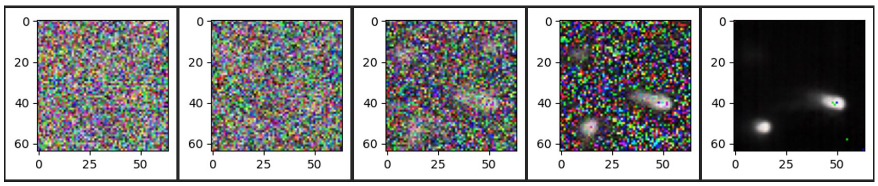

Figure 4: Sequential visualization of the denoising process using the diffusion model. The images depict the model's progression from a state of random noise to a clear representation of the melt pool. This reverse process occurs over several iterations, showcasing the model's ability to progressively reconstruct the initial image from noise.

Figure 4 displays the gradual transition of an image from noise to a coherent melt pool representation. As the diffusion model iteratively performs the backward process, each frame reveals more of the underlying structure. By the final frame, the model has significantly reduced the noise, revealing distinct features of the melt pool. This transformation not only validates the model's design but also its practical efficacy in image synthesis for Powder Bed Fusion processes.

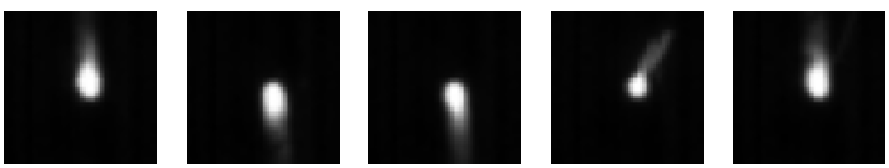

Figure 5: Original images from the dataset.

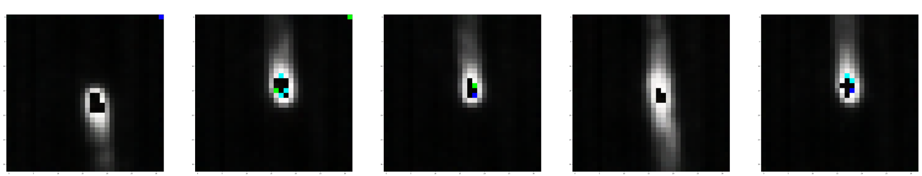

Figure 6: Generated images from the code.

We evaluate the similarities and differences between the original images from our dataset (Figure 5) and the images generated by our code (Figure 6). The comparison is aimed at assessing the quality of the generated images in terms of their resemblance to the original dataset.

- **Visual Similarity**: Upon visual inspection, the generated images exhibit a remarkable resemblance to the original images. Both sets of images maintain consistent themes and visual features that are indicative of the dataset's characteristics. The tonality, contrast, and general shapes present in the generated images align closely with those in the original set, suggesting that the generative model has successfully learned the underlying distribution of the dataset.

- **Fidelity of Generation**: The fidelity of the generated images is further evidenced by the preservation of intricate details that are unique to the originals. This includes the subtle gradations of shading and the precise replication of geometric patterns. Such details highlight the generative model's ability to capture and reproduce complex visual elements that define the dataset.

- **Differences and Limitations**: Despite the high degree of similarity, there are nuanced differences that can be observed. For instance, there is still some noise pixels still in the images, especially, the center of the melt pools. These variations, while minor, point to areas where the model might be further improved to enhance the quality of the image synthesis.

- **Inference**: Overall, the generated images stand as a testament to the capabilities of the generative model used in this study. While not without imperfections, the level of similarity achieved underscores the potential of diffusion models for tasks in medical imaging, art creation, etc. Future work may focus on refining the model to address the identified limitations and further improve the visual congruence with the original dataset.

7 Future Direction

7.1 Data-driven Image Generation

In the context of Powder Bed Fusion (PBF), controlling the quality of the melt pool is paramount. Traditionally, this control relies heavily on manually adjusting PBF parameters. However, by integrating data-driven techniques with the generative capabilities of Denoising Diffusion Probabilistic Models (DDPM), we can predict future melt pool images directly from process parameters.

Approach:

- **Data Integration**: Utilize existing datasets where each melt pool image is associated with specific PBF parameters stored in CSV files. These parameters, such as laser power, scan speed, and layer thickness, directly influence the melt pool morphology.

- **Model Training**: Adapt the DDPM to train on both images and corresponding parameters, enabling it to learn the complex mappings between the PBF parameters and the resultant melt pool characteristics. This training would leverage a modified DDPM architecture that integrates parameter inputs into the diffusion process.

Knowledge Alignment:

- **Concept**: Incorporate domain-specific knowledge into the diffusion process by enforcing physical constraints, such as energy conservation in the melt pool. This would ensure that generated images adhere to realistic physical laws, enhancing the reliability of the predictions.

- **Implementation**: Introduce a knowledge alignment network that adjusts the transition probabilities in the diffusion process based on the deviation from these constraints. By modifying the predicted mean of the denoising transition with an adjustment derived from the deviation of the generated image from the expected parameter-induced characteristics, the model can produce more accurate and physically plausible images.

Potential Benefits:

- This method would significantly improve the predictiveness and reliability of melt pool simulations, allowing for better anticipation of the final product qualities and more efficient parameter optimization.

7.2 Transformer-based Diffusion Model

Transformers have revolutionized many areas of machine learning by enabling models to handle data with complex dependencies. Applying Transformer architectures to diffusion models can potentially enhance their effectiveness for image generation tasks in additive manufacturing.

Advantages:

- **Attention Mechanisms**: Unlike U-Nets, which process data in a hierarchical manner, Transformers use self-attention mechanisms that weigh the importance of different parts of the input data, potentially providing a more nuanced understanding of complex relationships in melt pool dynamics.

- **Scalability and Flexibility**: Transformers can handle varying input sizes more flexibly, making them well-suited for integrating diverse data sources such as varying image resolutions or additional parameter data.

- **Improved Feature Capture**: The ability of Transformers to capture long-range dependencies can be particularly beneficial for understanding the intricate patterns and stability in melt pool formations, which are influenced by both local and global factors in the printing process.

Implementation Strategy:

- Transition from the current U-Net-based architecture to a Transformer-based model, ensuring that the model can still efficiently perform both the forward diffusion to introduce noise and the reverse process to reconstruct the image from noisy states.

- Implement and train the Transformer to evaluate its performance in comparison to the U-Net, particularly focusing on its ability to generate detailed and accurate melt pool images.

Potential Challenges:

- The complexity of Transformer models might require more computational resources, and careful tuning of model parameters will be necessary to achieve optimal performance without significant increases in training time.

8 Conclusion

This project demonstrated the capability of Denoising Diffusion Probabilistic Models (DDPM) to simulate and predict the dynamics of melt pools in Powder Bed Fusion (PBF) processes, significantly enhancing control and understanding of critical manufacturing parameters. By integrating real-world data with this advanced generative model, the research addressed prevalent challenges in PBF, such as the need for extensive trial-and-error experimentation, thereby promoting efficiency and reducing waste. The model's ability to generate high-fidelity images of melt pools suggests it can serve effectively as a decision-support tool in manufacturing settings, facilitating more precise adjustments to process parameters.

The DDPM exhibited robust performance in predicting and replicating the complex behaviors of melt pools, demonstrating potential beyond conventional image synthesis applications used in digital art or medical imaging. This research illustrates the versatility of diffusion models, offering a stable and scalable solution that surpasses traditional generative models like GANs in stability and output quality.

Looking ahead, integrating diffusion models with other AI technologies, such as reinforcement learning, could further revolutionize PBF and other manufacturing techniques by enabling systems that adapt dynamically to changing conditions. This would not only enhance the precision of manufacturing processes but also expand the practical applications of generative models in industrial settings, potentially leading to broader advancements in automated and intelligent manufacturing solutions.

9 References

References

[1] J. Ho, et al., "Denoising Diffusion Probabilistic Models," NeurIPS 2020, https://arxiv.org/abs/2006.11239.

[2] I. Kapelyukh, et al., "DALL-E-Bot: Introducing Web-Scale Diffusion Models to Robotics," IEEE Robotics and Automation Letters, 2023.

[3] "Diffusion Models: A Comprehensive Guide to Theory and Practice," AI Summer, https://theaisummer.com/diffusion-models/.

[4] J. Sander, "Understanding Diffusion Models: Bridging the Gap Between Generative Models and Energy-Based Models," 2022, https://sander.ai/2022/05/26/guidance.html.

[5] K. Piro, "Step-by-Step Visual Introduction to Diffusion Models," Medium, 2022, https://medium.com/@kemalpiro/step-by-step-visual-introduction-to-diffusion-models-235942d2f15c.

[6] "Introduction to Diffusion Models in Machine Learning," YouTube, Outlier, 2022, https://www.youtube.com/watch?v=TBCRlnwJtZU.

[7] "Diffusion Models Explained," YouTube, dtransposed, 2022, https://www.youtube.com/watch?v=S_il77Ttrmg.

[8] OpenAI, "ChatGPT: Optimizing Language Models for Dialogue," 2022. Available: https://www.openai.com.

[9] C. Doersch, "Tutorial on Variational Autoencoders," Carnegie Mellon / UC Berkeley, 2016, https://arxiv.org/abs/1606.05908.

[10] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," ICLR, 2014, https://arxiv.org/abs/1312.6114.

[11] I. Goodfellow, et al., "Generative Adversarial Nets," NeurIPS, 2014, https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf.