Using Diffuser generative models to synthetically generate steel microstructures

Zachary St. Thomas

Abstract:

Studying steel microstructures yields important insights regarding their mechanical characteristics. Within steel, microstructures transform based on a multitude of factors including chemical composition, transformation temperatures, and cooling rates. In particular, Martensite-austenite (MA) islands in bainitic steel appear as blocky structures with abstract shapes that are difficult to identify and differentiate from other types of microstructures. Machine learning models that are able to automatically and accurately detect these structures are a powerful tool in this regard, however the training process of the state-of-the-art machine learning models requires a large amount of high-quality data. As of recently, generative models are increasingly becoming capable of synthetically generating data of high enough quality to be used in training machine learning models. This project aims to implement Diffuser generative models for the purpose of generating synthetic examples of bainitic steel microstructures that would be capable for future research use, since obtaining authentic samples can prove expensive and time consuming, requiring the usage of powerful microscopes and high-definition imaging technology, something smaller research efforts may not be capable of. Ideally, the increased accessibility of unique, high-quality data will allow for more research to be done on detecting and analyzing MA islands and similar work can be done for other microstructures that appear in materials. The dataset for this project can be found here.

Introduction:

Bainitic steels have many uses in civil engineering such as production of rails, pipelines, and other forgings, as well as in the automotive industry to manufacture steering knuckles or chassis components. While research is ongoing into the origins and effects of MA islands in steel, it is hypothesized that their presence decreases the impact energy of the steel, making it less desirable than steel without the islands. Material scientists have been working on the identification, explanation, and replication of those structures. Understanding the mechanisms that lead to the formation of those specific morphologies is necessary to describe their impact on the mechanical properties. However, different sizes, shapes, and structures in bainite cause a manifold interpretation based on a subjective feature description by the expert who currently examines the microscopic images manually. This requirement to manually annotate data is a substantial expense in the scientific community, and while machine learning models have proven substantially suited for this annotation, the current lack of training data has left their performance still somewhat unreliable, especially for applications of scientific study. While there are certainly concerns of over-fitting to a subset of data, where novel images from generated data can be reverse-engineered to their training data origins, for a more narrow application of generative models to simply generate working training data, in many such instances this concern is less relevant.

Related Work:

DiffuGen: Training Generative Models with Diffusion Guidance, is a novel approach to training generative models using diffusion processes, involving using diffusion models as guidance during the generative process of a GAN (Generative adversarial network), which can help in creating more detailed and accurate outputs compared to traditional GANs alone. DiffuGen's approach of ensuring the stability of the training process by leveraging the gradual denoising process inherent in diffusion models, DiffuGen managed to mitigate some of the common challenges faced by conventional GANs such as training instability and mode collapse. Their framework was also made to be highly flexible, allowing for easy adaptation and augmentation to a wide variety of applications in generative modeling and data mining. The paper proposing DiffuGen included its architecture alongside experiments and examples that demonstrate its high-quality images. These experiments also compare the performance of DiffuGen against other state-of-the-art generative models, showing its competitiveness and potential advantages.

UniformGAN is another breakthrough work from the conference paper *Overcoming Challenges of Synthetic Data Generation* which addressed the limited use cases of diffusion models and the ability to reverse engineer the training data of a diffusion model from its output. It accomplishes this by utilizing a novel loss function for GANs, as well as SeLU rather than ReLU for an activation function. They pre-process input datasets to transpose each column into a uniform distribution, which allows for a wider application area for UniformGan. Additionally, they add privacy to the input dataset by adding noise to the discriminator during training, preventing easy re-engineering. The paper includes the architecture of UniformGan as well as a command-line tool to generate and evaluate synthetic datasets on numerous metrics.

Data:

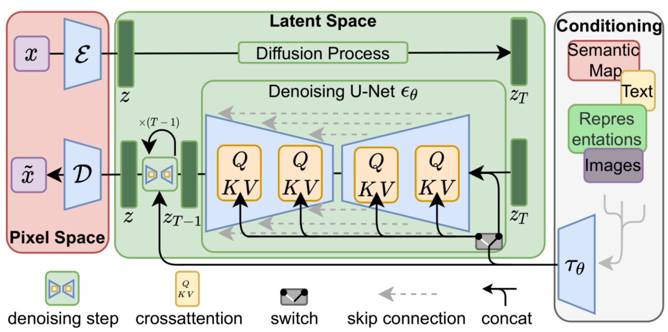
The initial dataset consists of 1,705 electron microscopy image scans, with 8,909 annotated polygons over the images surrounding the MA islands within the images. The MA images have their centers marked as points of interest (POIs) by the annotators, representing their belief of a MA island at that point. There were three annotators, and MA islands have additional markers for if two or three annotators believed a specific area to be a POI. These POIs were used as a central point for the annotators to draw their polygons around the MA islands and marked a defining aspect of the island that gave the annotator the belief that it was an MA island. The images are standardized to the same size, available in png format as well as tiff format. Since the structures are so highly similar at the microscopic level, the images share many features with one another and are thus prime to be used in image segmentation and generation applications.

The dataset also included two notebooks which contained Python code to both annotate the images with polygons surrounding the MA islands, as well as code for a baseline model which would be used to evaluate the performance of any future computer vision models. Both of these notebooks, the polygon notebook and the model notebook, were very useful to load in the data for the purposes of image segmentation and data generation.

Methods:

I used their model to analyze both the original and synthetic dataset. I also used an image to image Diffuser model to generate similar images and generate labels for the MA islands in the images it generates. I both used a pre-trained Diffusion model for this, as well as a U-net Diffusion model I had trained myself, in the interest of comparing the quality of the produced data. The labeling is a considerable challenge, since usually a team of experts will manually review and label the MA deposits, and a model that generates similar images will not intuitively know what would constitute MA deposits and what would not. Using their existing baseline model, however, is enough to ensure that such deposits are at least recognizable. However, to verify the integrity of any produced images, experts in Bainitic steels and Martensite-austenite microstructures would be required.

Figure 1: Diffusion U-net architecture.



Experiments:

To evaluate the performance of the synthetic data, I compared productions from the Diffusion models to the original data present in the dataset. I made this comparison for both the pre-trained Diffusion model and the one trained exclusively on the steel microstructures dataset. Through these comparisons, the images produced by the model trained exclusively on the steel proved to have more accurate re-creations of the training dataset. Images are included below for comparison.

Figure 2: Collage of images generated from the diffusion model trained exclusively on the dataset.

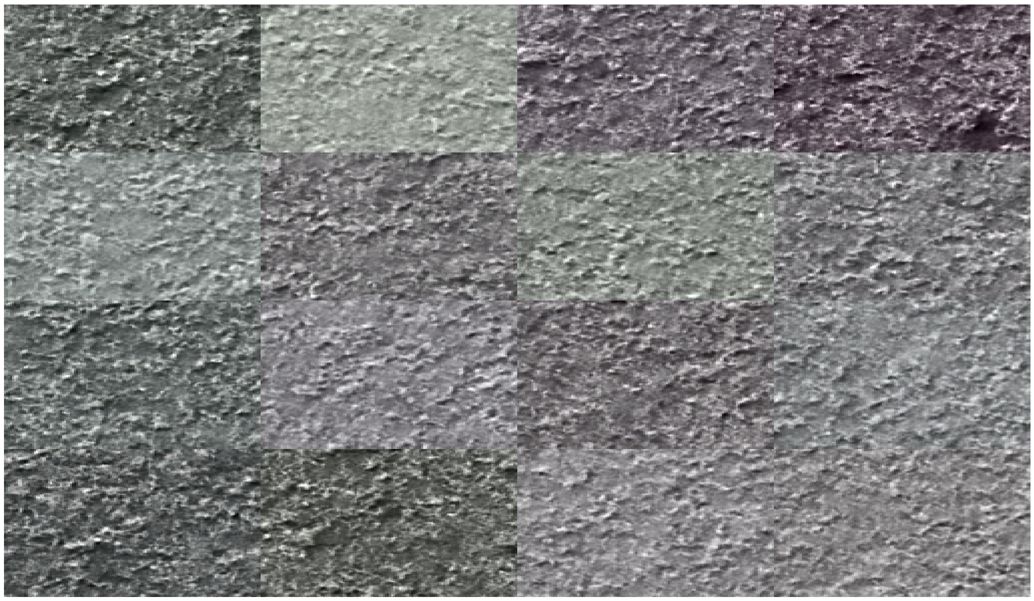
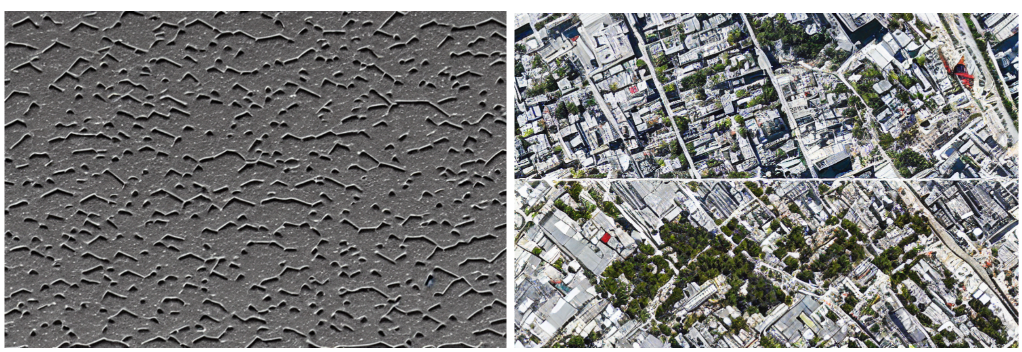


Figure 3: Two images containing two separate examples of images generated from the pre-trained model.



Included with this report will be directories of generated images, as well as the notebooks I used to generate all the images.

Conclusion:

The experiment results from the Diffusion models show that there is some potential in generating training data for various image segmentation applications, specifically from a custom-trained diffusion model. Though the specific application field of expert-annotated steel microstructures is a highly difficult one, and ultimately the data generated can only be verified by experts, the similarities were strong enough from a model trained exclusively on the images to prove capable of generating training data images for image segmentation tasks.

References:

1. Carlini, Nicholas, et al. “Extracting Training Data from Diffusion Models.” , 30 Jan. 2023.
2. Shenoda, Michael, and Edward Kim. “Diffugen: Adaptable Approach for Generating Labeled Image Datasets Using Stable Diffusion Models.” , 1 Sept. 2023.
3. Samuel, Dvir, et al. Generating Images of Rare Concepts Using Pre-Trained Diffusion Models, 27 Dec. 2023.
4. Yu, Ying, et al. “Techniques and Challenges of Image Segmentation: A Review.” *MDPI*, Multidisciplinary Digital Publishing Institute, 2 Mar. 2023.
5. Fang, Kevin, et al. *Overcoming Challenges of Synthetic Data Generation*, 26 Jan. 2023.
6. Nguyen, Quang Ho, et al. “Dataset Diffusion: Diffusion-Based Synthetic Data Generation For...” *OpenReview*, 2 Nov. 2023.