



Parameters, Properties, and Process: Conditional Neural Generation of Realistic SEM Imagery Toward ML-Assisted Advanced Manufacturing

Scott Howland¹ · Lara Kassab² · Keerti Kappagantula¹ · Henry Kvinge² · Tegan Emerson²

Received: 7 September 2022 / Accepted: 26 October 2022 / Published online: 24 January 2023
© Battelle Memorial Institute 2023

Abstract

The research and development cycle of advanced manufacturing processes traditionally requires a large investment of time and resources. Experiments can be expensive and are hence conducted on relatively small scales. This poses problems for typically data-hungry machine learning tools which could otherwise expedite the development cycle. We build upon prior work by applying conditional generative adversarial networks (GANs) to scanning electron microscope (SEM) imagery from an emerging advanced manufacturing process, shear-assisted processing and extrusion (ShAPE). We generate realistic images *conditioned* on temper and either experimental parameters or material properties. In doing so, we are able to integrate machine learning into the development cycle, by allowing a user to immediately visualize the microstructure that would arise from particular process parameters or properties. This work forms a technical backbone for a fundamentally new approach for understanding manufacturing processes in the absence of first-principle models. By characterizing microstructure from a topological perspective, we are able to evaluate our models' ability to capture the breadth and diversity of experimental scanning electron microscope (SEM) samples. Our method is successful in capturing the visual and general microstructural features arising from the considered process, with analysis highlighting directions to further improve the topological realism of our synthetic imagery.

Keywords Deep learning · Scanning electron microscopy · Conditional image generation · Generative adversarial networks

Introduction

We live in an age of unprecedented technological growth. This growth has changed the daily lives of people across the globe. Changes like more compact devices, improved battery life, and faster processing are accompanied by additional methods of communication and access to/sharing of information. The former are underpinned by exploiting new materials and advanced manufacturing technologies, while the latter come at the hand of increasingly sophisticated artificial intelligence (AI). Although there have been incremental advances in leveraging AI for specific analysis tasks in advanced manufacturing, there are no established, generalizable frameworks for accelerating research and development across material systems and manufacturing processes.

Because of the physical regimes in which they operate, advanced manufacturing processes are supported by nascent first-principle simulation capabilities instead of the

✉ Scott Howland
scott.howland@pnnl.gov

Lara Kassab
lara.kassab@pnnl.gov

Keerti Kappagantula
keertisahithi.kappagantula@pnnl.gov

Henry Kvinge
henry.kvinge@pnnl.gov

Tegan Emerson
tegan.emerson@pnnl.gov

¹ Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99354, USA

² Pacific Northwest National Laboratory, 1100 Dexter Ave N, Seattle, WA 98109, USA

more conventional or established approaches due to their cutting-edge nature. Owing to the ongoing emergence of such simulation or physics-based models for advanced manufacturing processes, practitioners are often forced to rely on their intuition and a small body of data when designing experiments which may or may not probe synthesis regimes that ultimately result in desired material microstructures and performance metrics. This can lead to a slower, more expensive research and development cycle and a delay of advanced manufacturing deployment for pilot and commercial-sale applications.

It is well known in the materials science and manufacturing fields that material microstructures play a central role in associating manufacturing process parameters used in synthesizing a component (or sample) and its performance. As such, microstructural features are crucial for guiding and interpreting manufacturing data. Of the multiple methods which enable determination of metal microstructures, scanning electron microscope (SEM) imaging is a popular approach for capturing information regarding important material features such as grain size distribution, precipitate morphology, and grain boundary density amongst others. However, SEM images must be analyzed to identify key features of interest, which requires domain knowledge and post-processing activities. All this makes SEM imaging a time and resource-intensive endeavor. Accordingly, there is great interest in reducing the number of SEM images that have to be obtained for a developmental process while also decreasing the cost of associated post-processing and analysis efforts.

There are several models available for predicting the microstructures of materials manufactured in a specific process parameter regime or identify the microstructures of the materials demonstrating a specific combination of performance metrics using first-principles approaches for conventional manufacturing approaches. However, as discussed above, such models are readily available to generate microstructures corresponding to either specific process conditions or final performance for advanced manufacturing. More recently, deep learning (DL) has found various applications to interpreting and understanding SEM images in the materials science and manufacturing applications, such as automatic classification of images [1–3] and segmentation of images to identify different regions of interest [4]. Recently, DL methods have been used to generate SEM images of different materials [5] and more [6]; however, it is important to note that in most of these works, the DL approach deals with images in isolation and is not explicitly informed by the manufacturing technique. It is well understood that microstructural features strongly depend on the manufacturing conditions used to produce them. Therefore, while it is a great advancement to use DL for generating SEM images, reducing the cost of associated research and

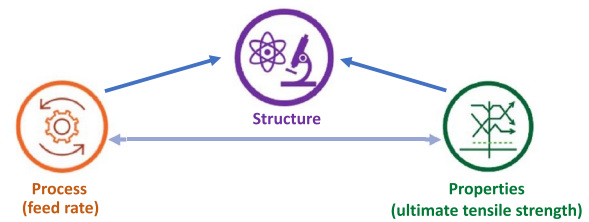


Fig. 1 Process parameter–microstructure–material property triangle. In this diagram, we identify the process parameter and resulting material property we consider for ShAPE manufacturing, namely “feed rate” and “ultimate tensile strength”, respectively

development activities, we also note that these prior works are unable to incorporate the valuable, process-dependent information necessary to expedite the development-validation cycle. Subsequently, there is a critical need to generate SEM images conditioned on specific manufacturing parameters as the next wave of DL development for materials and manufacturing image analysis. Incorporating a conditional component into SEM image generation enables the production of synthetic SEM imagery which conditionally depends on either manufacturing process parameters or target material properties as illustrated in Fig. 1. While conditional image generation models have been widely used, most DL techniques are data-hungry, which presents problems when applied to domains like materials science and advanced manufacturing which suffer precisely from scarce data. Therefore, it is essential for any conditional SEM image generation to be feasible even when trained on small datasets.

In this paper, we present a rationale and approach for addressing these challenges in applying DL for SEM imagery given limited training data. We demonstrate an ability to produce realistic microstructures in synthetic images and provide methods to quantify consistency with experimental SEM images through the use of topological feature extraction. This work takes a critical step toward leveraging machine learning to help accelerate advanced manufacturing research and development in light of developing first-principles simulations. We develop generative models trained on SEM images of aluminum alloy AA7075 tubes manufactured via the Shear-Assisted Processing and Extrusion (ShAPE) technology [7, 8]. ShAPE is an emerging advanced manufacturing process that synthesizes rods, bars, tubes, and wires [9–12] of different cross-sectional areas and shapes from metallic (pure metals, alloys) feedstock in various forms such as powders, chips, films, discs, and solid billets [13–15]. ShAPE-synthesized parts demonstrate unique microstructures with minimal porosity and never-before-seen performance. Several publications are available describing the synthesis and characterization of ShAPE samples made from aluminum, magnesium, copper, and steel, among others [16, 17]. ShAPE demonstrated

enhanced performance in bulk-scale components, making their scale-up pathways relatively viable for industry. Therefore, there is an urgent need to develop models which can associate ShAPE process parameters with resulting microstructures in order to reduce research and development time and deployment delays for ShAPE at an industrial scale.

Background and Related Work

Manufacturing with Shear-Assisted Processing and Extrusion (ShAPE)

This work is focused on generating SEM microstructures of aluminum alloy AA7075 tubes manufactured using ShAPE [8]. During ShAPE, a rotating die impinges on a stationary billet placed within an extrusion container equipped with a coaxial mandrel. At the interface between the billet and the die, the billet is heated and plasticized by both the shear forces applied and by the resulting frictional heat. As the die moves into the plasticized billet material, the material emerges from a cavity in the die to form a tube extrudate. ShAPE process parameters are comprised of data streams such as tool rotation rate and tool traverse rate (feed rate), which result in specific extrusion temperatures, forces, torques, and power.

The data used in this study were developed by [7], which resulted in manufactured AA7075 tubes. The authors then obtained tube coupons and different locations and subjected them to T5 and T6 heat treatment before finally testing them to determine their ultimate tensile strength (UTS), yield strength (YS) and elongation. The coupons were also imaged using a scanning electron microscope to obtain the fore-scatter and back-scatter images of the microstructure of the samples. Of the several process–microstructure–property

data streams available from the original study, in this work we narrowed our scope to consider a single ShAPE processing parameter, the feed rate, a single resulting material property, the UTS, and the back-scatter modality of the SEM images. We also account for the post-ShAPE heat treatment (T5, T6 tempers) as it can strongly influence material properties for ShAPE materials and those from other processes more broadly. We illustrate the interplay of process parameters, material properties, and microstructures from SEM back-scatter images in Fig. 1 by generating conditioned SEM images corresponding to either specific feed rates and temper at which the samples were manufactured or UTS (Fig. 2).

Generative Adversarial Networks

Generative adversarial networks (GANs) [18] are designed to model a training data distribution by way of an adversarial game played by two neural networks: a generator and a discriminator network. The generator takes as input a noise vector, typically sampled from a standard normal distribution, and uses this entropy source to generate a unique data sample—for our purposes, an image. The discriminator takes as input a data sample and predicts whether it is a real sample from the training set, or whether it is a fake sample produced by the generator. The networks are adversarial in that they are optimized to fool one another, with the goal of producing a generator which can sample data indistinguishable from the original training distribution.

Since their introduction, GANs have undergone multiple extensions and improvements: some aimed to stabilize their training dynamics, others to improve their generation quality, and others still to allow GANs to incorporate new sources of domain information. Conditional GANs (CGANs) [19] provide side channel information, typically class labels, to both the generator and discriminator networks. This facilitates

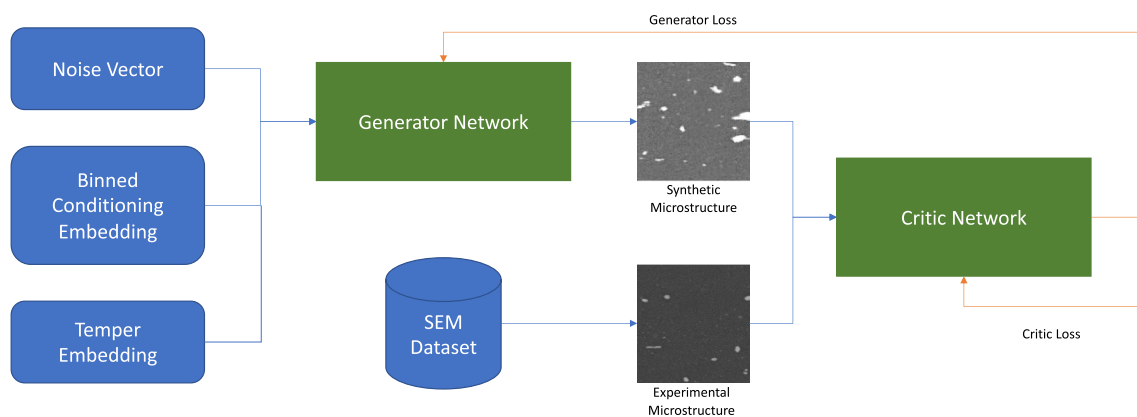


Fig. 2 A high-level overview of our GAN model's component and data flow. Blue elements are associated with input data, while green elements are trainable neural network components. Blue arrows constitute the model's forward pass, whereas orange arrows represent backpropagation updates to the model's parameters

sampling from sub-distributions of the original training data. This approach was improved upon by auxiliary classifier GANs (AC-GANs) [20], which only provide side channel information to the generator and task the discriminator with learning to determine both authenticity *and* the side channel information of a given sample.

Another avenue for GAN improvement arrived with the introduction of the Wasserstein GAN [21], which introduced a new loss formulation based on the Wasserstein distance [22] to improve training stability. This formulation changed the discriminator to produce an authenticity *score* rather than a simple real-fake classification, lending these particular discriminators the title of “critics”. Later work in this vein introduced a gradient penalty to the WGAN’s loss function (WGAN-GP) [23], improving training stability yet further.

Microstructure Topological Feature Analysis

Persistent homology is a popular tool from topological data analysis used to study the shape of data [24]. In particular, *sublevel set persistent homology* is a technique frequently applied to grayscale image data to study the variation in intensity of patches (neighboring pixels) in an image.¹ In this context, a $m \times n$ persistence image (PI) is often used to represent the persistence of topological features of the data across scaling [25]. It summarizes the creation time (or *birth*) on the horizontal axis and persistence time on the vertical axis as topological features.

Recent work [26] has demonstrated that persistence images can act as powerful and robust feature descriptors for microstructures like those collected from ShAPE AA7075 tubes—specifically microstructures well characterized by precipitate intensity/contrast and precipitate density/distribution. Given the interpretability, generalizability, and noise robustness provided by these methods, we leverage them in evaluating the fidelity between experimental and synthetic SEM microstructures. By comparing PIs from experimental and synthetic SEM images, we are able to visualize and discuss similarities and differences between the two image types (real and generated) in a feature space that preserves domain knowledge [26]. The relationships between PIs are visualized by using principal component analysis (PCA) to perform dimensionality reduction so that both groups can be visualized together in 3-dimensions.

Experimental Methods

SEM Image Dataset

We took as our GANs’ input data a collection of SEM images of 32 AA7075 ShAPE tubes manufactured using 32 different process conditions followed by T5 and T6

heat treatments. The SEM microstructure images are large (2560×1920 pixels), grayscale, and were taken at 500× magnification. The most important feature of interest in the AA7075 ShAPE tube SEM images are the intermetallic precipitates which are seen as lighter particles with varying morphology and topology. However, 32 samples is far too little to train a deep learning model, but due to the relatively small precipitate size in these images and the lack of large salient precipitates we cropped each large SEM image into smaller chips of size 128×128. The chips were partially overlapped and yielded a training set of 437,000 images which we used in our experiments.

Our data preprocessing choices were enabled by the precipitate form and structure itself. In the ShAPE AA7075 tube microstructures, the critical microstructures can be observed over a fairly limited, local spatial extent. We arrived at our crop size by identifying the smallest chip that a human expert would be capable to effectively evaluating. This choice further depends on the magnification level. If, for another manufacturing process, the entire SEM image is needed for microstructure evaluation, it may be more practical to learn to generate descriptions or *features* of the SEM images rather than the entire image when only a small number of training samples are available. Future efforts will explore ways to numerically determine these parameter choices.

Even with a larger number of images to work with, we still have a very small number of overall experiments, which are the source of the experimental parameter and material property values we use for conditioning information—namely ShAPE feed rate and UTS. Initial experiments where we conditioned our models on normalized scalar values did not perform well. We suspect this was because the conditioning information was simply too sparse, and that the models overfit and were unable to extrapolate beyond this small experiment set.

As has been done elsewhere in the literature [27–30], we relaxed the desired regression problem into a classification problem by discretizing the feasible range of values into categorical variables. In doing so, we transformed the conditioning information (either process parameters values or property values) into three categories: “low”, “mid”, and “hi”. These bins divided the scalars observed in our experiments into lower, middle, and upper thirds. This is a coarse binning, but one that could still significantly accelerate manufacturing research if predicted accurately by providing guidance about possible SEM microstructures that can be obtained when a samples is manufactured in a specific regime that demonstrates UTS above or below a certain user-specified number.

Additionally, nine of the ShAPE experiments produced extrudates which did not undergo further T5 or T6 tempering. Since UTS was not available for the ShAPE tubes

¹ In our analysis, we use one-dimensional homology H_1 .

that were not heat treated further, we only include these “as extruded” experimental microstructures when learning the effects of ShAPE feed rate on resulting microstructures. This temper imbalance, combined with imbalances between our “low”, “mid”, and “hi” conditioning labels, gave us the per-setting experiment counts depicted in Table 1.

GAN Architecture and Training

For our generative models, we used Wasserstein GANs with gradient penalty [23] and an auxiliary classifier, abbreviated ACWGAN-GP. Its generator module takes as input the concatenation of two inputs: a 100-dimensional noise vector sampled from a standard normal distribution, as well as learned, 20-dimensional dense embedding vectors for two tempering conditions (corresponding to T5 and T6 treatments) and similarly learned 20-dimensional “low”, “mid”, and “hi” label embeddings for either a manufacturing experimental parameter or a material property. From these inputs, which we concatenate along the feature dimension, the generator produces a grayscale image. The GAN’s critic module takes as input a grayscale

image and produces separate scores for: the WGAN critic score (scalar), temper classification (two logits), and the correct binning label for the experimental parameter or material property (three logits). We trained two GANs in our experiments: one conditioned on feed rate and one conditioned on desired UTS.

We trained all GANs for 400 epochs over our training set of SEM crops, with a batch size of 64 and the AdamW optimizer [31]. During training, we upweighed each model’s gradient penalty by a factor of 10 and the critic’s classification losses by a factor of 5. We found these values gave an effective training speed as well as diversity and quality of synthetic images.

Our approach is similar to that of [5], with differences emerging from our data and our particular framing of the conditional generation problem. The ShAPE data afforded us a significantly greater number of training samples (437,000 vs approx. 7,000), each of which has a “simpler” visual structure (ex. Figures 3 and 4) than samples from the Ultra High Carbon Steel DataBase (UHCSDB) [32]. Our work examined two tempering conditions (T5 and T6) instead of five, though we explored joint conditioning over both temper and parameter or material property conditions. That our work explores conditional generation over process parameters and material properties is novel in itself.

Table 1 The number of experiments within our “low”, “mid”, and “hi” label binnings for feed rate and ultimate tensile strength (UTS) conditions for both T5 and T6 temper treatments

Temper	Feed rate label counts	Ultimate tensile strength label counts
T5	9 / 8 / 15	11 / 7 / 3
T6	9 / 8 / 15	12 / 5 / 4

As a process parameter selected before an experiment begins, feed rate is independent of tempering. Bin values correspond, respectively, to lower, middle, and upper thirds of observed values

Results and Analysis

Synthetic Image Quality

Figures 3 and 4 show synthetic images conditioned on T5 and T6 temper conditions jointly with “low”, “mid”, and “hi” ultimate tensile strength (UTS) or feed rate, respectively. We pair the synthetic images with experimental ones

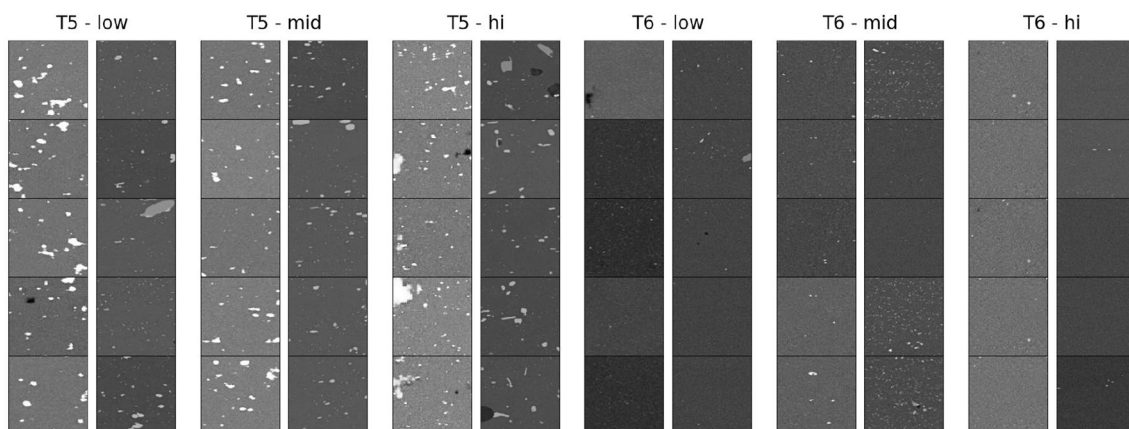


Fig. 3 Paired columns of synthetic (left) and experimental (right) ultimate tensile strength-conditioned SEM images under T5 and T6 temper settings

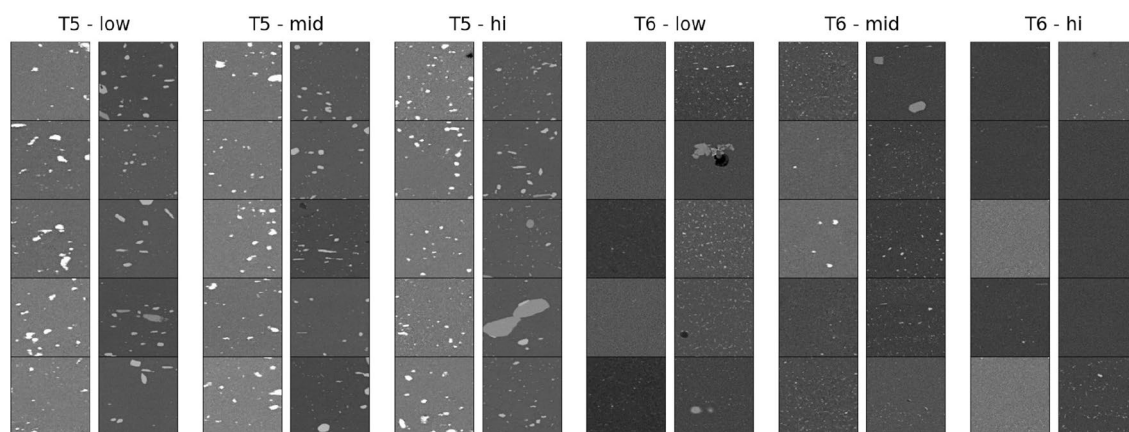


Fig. 4 Paired columns of synthetic (left) and experimental (right) feed rate-conditioned SEM images under T5 and T6 temper settings

for comparison. Among some notable differences between the experimental and synthetic images, we see that exceptionally large precipitates appear more artifacted and less frequently in synthetic images than in experimental ones, especially under the T6 condition.

Despite the heavy imbalance of binned labels across experiments shown in Table 1, we saw synthetic T5 and T6 samples consistently display high and low degrees of sample diversity, respectively, regardless of the number of experiments associated with a given label bin. We suspected this was because T6-tempered microstructures tend to have far fewer salient precipitates against a background of tiny, visually near-random precipitates. This produces less variation throughout an uncropped SEM image, and by extension less variation among our training crops. This hypothesis is also consistent with the lack of mode collapse (a common issue encountered when training GANs wherein the model learns to produce a single plausible output at the expense of output diversity) in the T5 setting, where the presence of many larger, irregular precipitates means crops from within a single SEM image will be highly diverse. Additionally, the artifacting of large precipitates in the T5 condition could benefit from differentiable data augmentation [33] as a way to upsample rare phenomenon in our experimental data. Incorporating recent GAN techniques targeting efficient use of unbinned, scalar conditioning labels [34] could further improve performance by removing the need for a discretely-conditioned, categorical latent space in favor of more continuous ones, able to better leverage what is presently unwieldy conditioning information. Recent GAN regularization techniques based on consistency [35] and separating the discrimination and classification tasks [36] could supplement these approaches more generally.

While visual quality is an important indicator of our GANs' performances, it does not tell the full story. In order to be useful for experimental design and property analysis,

our models must produce images which are physically meaningful, not just visually plausible. Our remaining evaluations are focused on understanding the degree to which our GANs' image distributions align with our experimental ShAPE data.

Topological Fidelity Experiments

We average the 10×10 PIs of experimental and synthetic images: Fig. 5 shows average PIs derived from experimental samples as well as all of our GANs. Briefly, each PI is constructed by gradually increasing a threshold which is used to capture spatially contiguous pixels with intensity up to the threshold. As the threshold increases, more pixel clusters (precipitates in our case) will fall below the threshold and be "born". This is captured by the horizontal axis: pixels further to the right capture pixel clusters that are "born" under a later threshold—in our case, it captures brighter precipitates. The PIs also capture how long a given cluster "lives" before eventually merging with other clusters as the threshold continues to increase. This merging is based on the proximity and relative intensity of these neighboring clusters and so captures their spatial distribution. Pixels higher along the vertical axis represent precipitates that have longer "persistence". For example, we would expect a microstructure with a blend of low and high-brightness precipitates to have a PI with clusters of pixels on both the left and right; we would also expect the PI of a microstructure with a few large, distant, bright precipitates to have more pixels concentrated toward the top of the image than a microstructure with many small precipitates evenly sprinkled throughout, even if the brightness of these precipitates were the same in both cases.

At a high level, we can see that in our synthetic and experimental images, the general shape of the persistence pattern is preserved; however, the synthetic images do

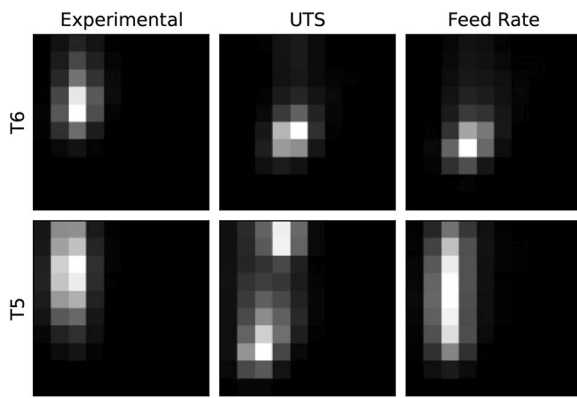


Fig. 5 Average persistence images (PIs) over experimental (top) and synthetic (bottom) ShAPE SEM chips. The PIs we use detect the number of “holes” in an image, a statistic that has been shown to aligned with scientifically salient features. PIs that are similar indicate that the number and scale of “holes” between two images is similar. We compare across synthetic imagery conditioned on either feed rate or ultimate tensile strength (UTS) and across T5 and T6 temper conditions

not have precise agreement with the experimental image distributions. Interestingly, we see that our synthetic images produce an overly concentrated persistence pattern in the T5 case, but an overly diffuse persistence pattern in the T6 case. Since these are *average* persistence images, we conclude that our distribution of synthetic T5 images exhibit insufficient levels of structural variation (mild mode collapse), whereas the T6 distribution exhibits unrealistically *high* levels of variation. The overly “tall” T6 persistence patterns are consistent with our observation that our GANs are able to rely too heavily on producing a blanket of tiny, noisy precipitates compared to experimental references. Partial mode collapse is also visible in Figs. 3 and 4 in the form of synthetic samples of a consistent, high average brightness. Some experimental samples have a comparable degree of brightness but the synthetic image distribution fails to capture the experimental diversity in both precipitate and background brightness. This problem might be alleviated by denoising the backgrounds of experimental images which would free the generator and discriminator from needing to model blankets of noise, especially in the T6 setting. Such denoising would not introduce additional training signal, but reducing the degree of confounding or task-irrelevant visual information could aid the GAN optimization process.

At the same time, the synthetic persistence patterns are still close to their experimental counterparts. The exception would be the PI for the synthetic images corresponding to T6 condition conditioned on UTS, which exhibits an oddly bimodal persistence pattern indicative of partial mode collapse. These results are consistent with GANs which are trained to mimic the visual distribution without access to

topological or physical regularization, which could better align them with experimental SEM data.

Equipped with PIs as features, we perform principal component analysis (PCA) [37] to reduce the dimensions of experimental imagery so that it can be visualized. We can then inspect the alignment between our experimental and synthetic data. Figure 6 visualizes the projection of our synthetic and experimental PI images onto the experimentally-fit components for either feed rate or ultimate tensile strength (UTS)-based conditioning. We can see in both cases that our synthetic PIs do not overlap well with the experimental PIs for both T5 and T6 temper conditions; indeed, they appear nearly orthogonal in component space. We also note that there is exceptionally less agreement between T5 and T6 projections for the synthetic, feed rate-conditioning case than all other settings. This is consistent with the average PIs in Fig. 5 and with the visual samples in Figs. 4 and 3. The average PIs reflect a similar dynamic due to the feed rate-conditioned T6 PIs having an unrealistically high spread of intensities compared to all other settings, whereas the image samples show that the feed rate-conditioned T6 model is poor at capturing the larger precipitates in the experimental data—even compared to our UTS-conditioned model.

These PCA results provide evidence that our model generators learn a manifold which is *visually* similar to the ground truth data distribution, at least in the T5 setting, but which is *topologically* quite distinct. This is counter to intuition around GANs where in theory the discriminator will pressure the generator into matching the true data distribution, in all its aspects, over time. However, in general the history of GAN research shows the myriad ways in which this process can fail to live up to its theoretical potential: unstable optimization dynamics, imbalanced generator-discriminator training, and other difficulties can produce unsuccessful or only partial alignment between the generator and ground truth data distributions. While issues of imbalanced generator-critic steps are largely accounted for by our Wasserstein GAN architecture, we experimented with more powerful critics to verify that this issue was not caused by underparameterization. Specifically, we replace our several-layer convolutional network with a ResNet-18 [38] either randomly initialized or pretrained on ImageNet [39] and see *degraded* performance compared to our initial, smaller architecture. We hypothesize that this discrepancy is not due to lack of model capacity but rather due to challenges in the optimization process: the generator can only be pressured into adopting topologically realistic data if the critic itself can distinguish disparate topological features. Incorporate explicit topological regularization, or a second small critic network providing feedback on derived persistence images, would be a useful avenue for future work to enforce a more physically plausible synthetic image distribution.

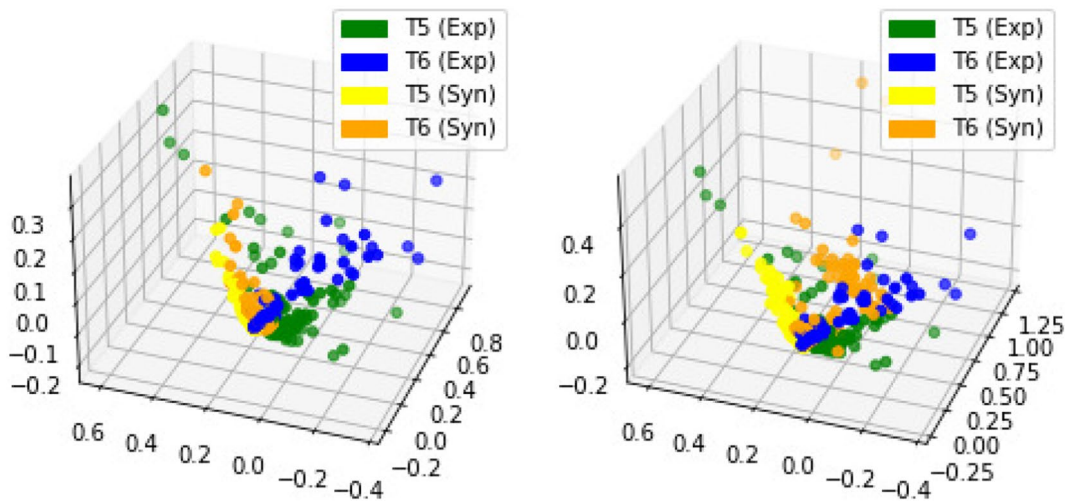


Fig. 6 We project persistence images into a 3D space using principal component analysis to reduce the dimensions of our experimental data, with synthetic images sampled from our models conditioned on either ultimate tensile strength (left) or feed rate (right)

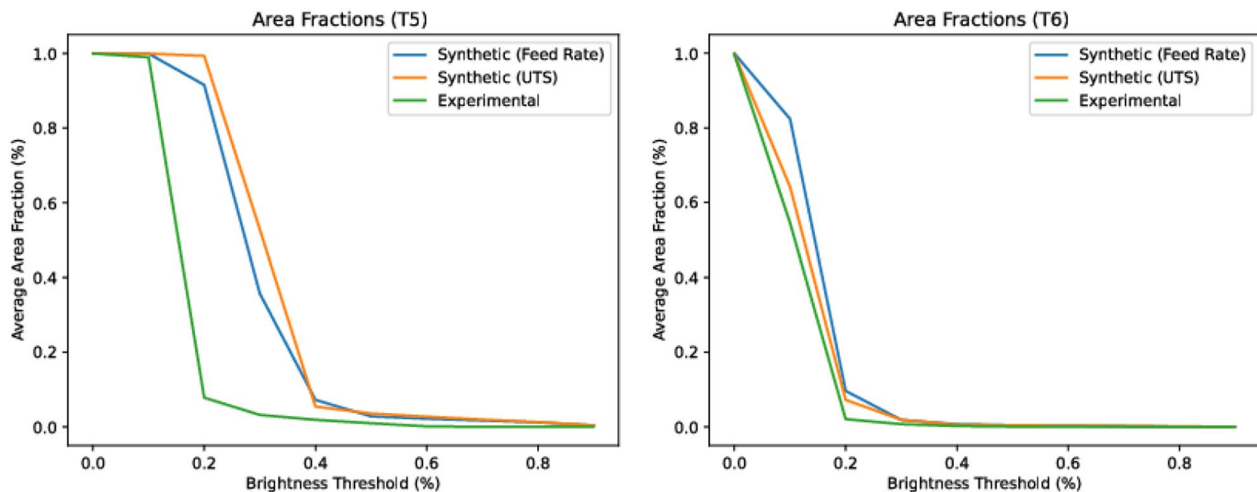


Fig. 7 Area fraction statistics for our synthetic and experimental data across T5 and T6 temper conditions. Fractions are calculated using 1,000 images randomly selected from each condition. We performed

mean centering and scaling to $[0, 1]$ within each condition for the sake of comparison

Finally, we examined the area fractions of bright pixels across synthetic and experimental microstructures; this is done because such area fractions are a common heuristic used in material science applications to evaluate precipitate density in microstructures. Because we were interested in comparing area fractions across conditions in order to ascertain the prevalence of bright precipitates against the microstructures' darker backgrounds, we performed feature scaling within each condition such that images share a center mean and pixel intensities lie in the range $[0, 1]$. In order to measure the area fraction of a given image, we then set a threshold $t \in [0, 1]$ and calculated the percentage of pixels

with intensity $i \geq t$. We used ten evenly spaced thresholds in $[0, 1]$, with the results captured by Fig. 7.

The results show that the variation in mean area fraction between synthetic and experimental T5 microstructures is less than 10 percent. This implies that in the generated SEM images, the precipitates occupy an area fraction that roughly coincides with that of the experimental images. This is not the case for T6 microstructures, where we observe that experimental datasets have consistently higher area coverage of darker to mid-tone precipitates. This is consistent with both our visual observations in Figs. 3 and 4 and in the persistence images in Fig. 5. We suspect this phenomenon is caused by the lower precipitate density in our experimental

T6 data: when each image has low precipitate density, our models are able to minimize their loss functions by producing disproportionately large amounts of background noise in all cases. Regularizing our models by including loss terms for persistence image patterns or for area fraction scores could alleviate this problem and would be productive avenues for future work.

These results indicate that, despite an intuitive visual plausibility, albeit one that does not fully capture the natural variation of experimental data, our models are learning a distribution over SEM images that is *topologically* distinct from experimental data. The generated images show precipitate morphology and distribution which result in similar area fractions as those observed in T5 experimental microstructures, though less so for T6 microstructures. Given the utility of topological features for characterizing SEM imagery [26], we pose improving this alignment as a useful avenue for future work. However, these limitations are unsurprising for a generative model trained on largely visual stimuli alongside a simple conditional embedding; the visual quality of our synthetic data indicates that it is possible to train generative deep learning models even on the relatively small data scale afforded by the advanced manufacturing domain, and our ability to pinpoint deficiencies in terms of latent space or topological feature alignment allows us to see actionable steps toward more physically realistic, deployment-ready development of this approach.

Conclusion

This work takes a critical first step toward a functional machine learning-accelerated advanced manufacturing experimental pipeline. We trained multiple conditional Wasserstein GANs (ACWGAN-GPs) on SEM microstructure image crops derived from AA7075 manufactured using the advanced ShAPE process. This is an advance over prior work which focuses on *unconditional* SEM generation for steels, marking a step toward a generative system that scientists can query to predict how either process parameters or properties impact a material microstructure. We observe that our synthetic images are visually plausible, though with some visual artifacting of rare precipitate phenomena. Additionally, we observe through topological methods of inquiry that our synthetic image distributions do not uniformly align with experimental SEM images. Specifically, we see dissimilarity between experimental and synthetic images in ways consistent with the small number of unique experiments present in most advanced manufacturing datasets. In future work, we propose exploring two avenues to address these limitations: differentiable data augmentation and recent developments in GAN regularization as a way to better leverage limited advanced manufacturing data and

to increase model sample efficiency, as well as topological and physical regularization to encourage the GANs to produce synthetic data which expresses even higher fidelity to experimental data distributions.

Acknowledgements This work was performed using the resources available at the Pacific Northwest National Laboratory (PNNL) and funded by the Mathematics for Artificial Reasoning in Science (MARS) Initiative as a Laboratory Directed Research and Development Project. KSK thanks Scott Whalen, Md. Reza-E-Rabby, Tianhao Wang, Scott Taysom, and Timothy Roosendaal for the ShAPE AA7075 tubes synthesis and property data. KSK also appreciates Woongjo Choi for their contributions to data arrangement; Tianhao Wang, Xiaolong Ma, and Alan Schemer-Kohn for developing the SEM images of the ShAPE AA7075 tubes; and Luke Gosink, Elizabeth Jurrus, and Sam Chatterjee for their support and advice on this project. PNNL is a multi-program national laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC05-76RL01830.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Azimi SM, Britz D, Engstler M, Fritz M, Mücklich F (2018) Advanced steel microstructural classification by deep learning methods. *Sci Rep* 8(1):1–14
2. Müller M, Britz D, Ulrich L, Staudt T, Mücklich F (2020) Classification of bainitic structures using textural parameters and machine learning techniques. *Metals*. <https://doi.org/10.3390/met10050630>
3. Tsutsui K, Terasaki H, Uto K, Maemura T, Hiramatsu S, Hayashi K, Moriguchi K, Morito S (2020) A methodology of steel microstructure recognition using sem images by machine learning based on textural analysis. *Mater Today Commun* 25:101514. <https://doi.org/10.1016/j.mtcomm.2020.101514>
4. Durmaz AR, Müller M, Lei B, Thomas A, Britz D, Holm EA, Eberl C, Mücklich F, Gumbsch P (2021) A deep learning approach for complex microstructure inference. *Nat Commun* 12(1):6272. <https://doi.org/10.1038/s41467-021-26565-5>
5. Iyer A, Dey B, Dasgupta A, Chen W, Chakraborty A (2019) A conditional generative model for predicting material microstructures from processing methods. arXiv preprint [arXiv:1910.02133](https://arxiv.org/abs/1910.02133)
6. Baskaran A, Kautz EJ, Chowdhary A, Ma W, Yenner B, Lewis DJ (2021) The adoption of image-driven machine learning for microstructure characterization and materials design: a perspective. *JOM* 73(11)

7. Whalen S, Olszta M, Reza-E-Rabby M, Roosendaal T, Wang T, Herling D, Taysom BS, Suffield S, Overman N (2021) High speed manufacturing of aluminum alloy 7075 tubing by shear assisted processing and extrusion (shape). *J Manuf Process* 71:699–710. <https://doi.org/10.1016/j.jmapro.2021.10.003>
8. Whalen S, Reza-E-Rabby M, Wang T, Ma X, Roosendaal T, Herling D, Overman N, Taysom BS (2021) Shear assisted processing and extrusion of aluminum alloy 7075 tubing at high speed. In: *Light metals 2021*. Springer, pp 277–280
9. Kalsar R, Ma X, Darsell J, Zhang D, Kappagantula K, Herling DR, Joshi VV (2022) Microstructure evolution, enhanced aging kinetics, and mechanical properties of aa7075 alloy after friction extrusion. *Mater Sci Eng A* 833:142575
10. Li X, Wang T, Ma X, Overman N, Whalen S, Herling D, Kappagantula K (2022) Manufacture aluminum alloy tube from powder with a single-step extrusion via shape. *J Manuf Process* 80:108–115
11. Li X, Zhou C, Overman N, Ma X, Canfield N, Kappagantula K, Schroth J, Grant G (2021) Copper carbon composite wire with a uniform carbon dispersion made by friction extrusion. *J Manuf Process* 65:397–406
12. Reza-E-Rabby M, Wang T, Canfield N, Roosendaal T, Taysom BS, Graff D, Herling D, Whalen S (2022) Effect of various post-extrusion tempering on performance of aa2024 tubes fabricated by shear assisted processing and extrusion. *CIRP J Manuf Sci Technol* 37:454–463
13. Darsell JT, Overman NR, Joshi VV, Whalen SA, Mathaudhu SN (2018) Shear assisted processing and extrusion (shape) of az91e flake: a study of tooling features and processing effects. *J Mater Eng Perform* 27(8):4150–4161
14. Taysom BS, Ma X, DiCiano M, Skrzek T, Whalen S, et al. (2022) Fabrication of aluminum alloy 6063 tubing from secondary scrap with shear assisted processing and extrusion. In: *Light metals 2022*. Springer, pp 294–300
15. Wang T, Gwalani B, Silverstein J, Darsell J, Jana S, Roosendaal T, Ortiz A, Daye W, Pelletiers T, Whalen S (2020) Microstructural assessment of a multiple-intermetallic-strengthened aluminum alloy produced from gas-atomized powder by hot extrusion and friction extrusion. *Materials* 13(23):5333
16. Jiang X, Whalen SA, Darsell JT, Mathaudhu S, Overman NR (2017) Friction consolidation of gas-atomized fesi powders for soft magnetic applications. *Mater Charact* 123:166–172
17. Whalen S, Overman N, Joshi V, Varga T, Graff D, Lavender C (2019) Magnesium alloy zk60 tubing made by shear assisted processing and extrusion (shape). *Mater Sci Eng A* 755:278–288
18. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. *Adv Neural Inf Process Syst* 27:548
19. Mirza M, Osindero S (2014) Conditional generative adversarial nets. *arXiv preprint [arXiv:1411.1784](https://arxiv.org/abs/1411.1784)*
20. Odena A, Olah C, Shlens J (2017) Conditional image synthesis with auxiliary classifier GANs. In: Precup D, Teh YW (eds) *Proceedings of the 34th international conference on machine learning*. *Proceedings of machine learning research*. PMLR, vol. 70, pp 2642–2651
21. Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. In: Precup D, Teh YW (eds) *Proceedings of the 34th international conference on machine learning*. *Proceedings of machine learning research*. PMLR, vol. 70, pp 214–223
22. Kantorovich LV (1939) The mathematical method of production planning and organization. *Manage Sci* 6(4):363–422
23. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC (2017) Improved training of wasserstein gans. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R (eds) *Advances in neural information processing systems*. Curran Associates, Inc, New York
24. Edelsbrunner H, Letscher D, Zomorodian A (2000) Topological persistence and simplification. In: *Proceedings 41st annual symposium on foundations of computer science*. IEEE, pp 454–463
25. Adams H, Emerson T, Kirby M, Neville R, Peterson C, Shipman P, Chepushtanova S, Hanson E, Motta F, Ziegelmeier L (2017) Persistence images: a stable vector representation of persistent homology. *J Mach Learn Res* 18:458
26. Emerson T, Kassab L, Howland S, Kvinge H, Kappagantula KS (2022) Toptemp: parsing precipitate structure from temper topology. In: *ICLR 2022 workshop on geometrical and topological representation learning*
27. Ahmad A, Khan SS, Kumar A (2018) Learning regression problems by using classifiers. *J Intell Fuzzy Syst* 35(1):945–955
28. Ammar Abbas S, Zisserman A (2019) A geometric approach to obtain a bird's eye view from an image. In: *Proceedings of the IEEE/CVF international conference on computer vision workshops*, pp 0–0
29. Truong L, Choin W, Wight C, Coda E, Emerson T, Kappagantula K, Kvinge H (2021) Differential property prediction: a machine learning approach to experimental design in advanced manufacturing. In: *AAAI 2022 workshop on AI for design and manufacturing (ADAM)*
30. Workman S, Zhai M, Jacobs N (2016) Horizon lines in the wild. *arXiv preprint [arXiv:1604.02129](https://arxiv.org/abs/1604.02129)*
31. Loshchilov I, Hutter F (2017) Decoupled weight decay regularization. *arXiv preprint [arXiv:1711.05101](https://arxiv.org/abs/1711.05101)*
32. DeCost BL, Hecht MD, Francis T, Weblar BA, Picard YN, Holm EA (2017) Uhcsdb: ultrahigh carbon steel micrograph database. *Integr Mater Manuf Innov* 6(2):197–205
33. Zhao S, Liu Z, Lin J, Zhu J-Y, Han S (2020) Differentiable augmentation for data-efficient gan training. In: Larochelle H, Ranzato M, Hadsell R, Balcan MF, Lin H (eds) *Advances in neural information processing systems*. Curran Associates Inc, New York
34. Zheng Y, Zhang Y, Zheng Z (2021) Continuous conditional generative adversarial networks (cGAN) with generator regularization. *arXiv preprint [arXiv:2103.14884](https://arxiv.org/abs/2103.14884)*
35. Zhao Z, Singh S, Lee H, Zhang Z, Odena A, Zhang H (2021) Improved consistency regularization for gans. *Proc AAAI Conf Artif Intell* 35(12):11033–11041
36. Li C, Xu K, Zhu J, Liu J, Zhang B (2021) Triple generative adversarial networks. *IEEE Trans Pattern Anal Mach Intell*. <https://doi.org/10.1109/TPAMI.2021.3127558>
37. Pearson K (1901) Liii on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine J Sci* 2(11):559–572
38. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
39. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al (2015) Imagenet large scale visual recognition challenge. *Int J Comput Vision* 115(3):211–252