Titanium 3D Microstructure for Physics-based Generative Models: A Dataset and Primer

Devendra K. Jangid* [1] **Neal R. Brodnik*** [2] **McLean P. Echlin*** [3] **Samantha H. Daly** [2] **Tresa M. Pollock** [3] **B.S. Manjunath** [1]

Abstract

When engineers design components, they rely on accurate property descriptions of the materials being used to predict performance. Most materials used for engineering applications are composed of an arrangement of atomic constituents into crystalline phases, which control the properties of that material. The crystal orientations embedded in this microstructural information differ from the information in conventional light optical images, and are critical for developing and designing materials for a range of applications. However, collecting microstructure information through experimental methods is expensive and time-consuming, especially when 3D information is needed. In order to model material properties under different material processing conditions (resulting in different microstructural arrangements), physics-based generative models are needed to create realistic synthetic microstructures. This research releases microstructural data of a titanium alloy, Ti-6Al-4V, and discusses their information modalities and the physics needed to be incorporated to enable the design of physics-based generative models for generating synthetic microstructures.

1. Introduction

Most materials used for engineering applications are composed of an arrangement of elemental constituents into crystalline phases, which control the properties of that material. The arrangement of these crystals (also known as grains) is referred to as the microstructure, and is described using metrics that capture the size, arrangement, connectivity, and crystallographic orientation of the grains. Efficiently obtaining sufficient microstructure information to predict material properties is crucial to the development of new technologies, especially under extreme environments. Voxel-based 3D microstructure information is especially valuable for understanding material behavior (Rowenhorst et al., 2010; Kelly et al., 2016; Hémery et al., 2019; Echlin et al., 2021; DeMott et al., 2021; 2020), but the cost of 3D reconstruction has limited many investigations to 2D characterization. Even when considering just 2D information, a large number of image sets have been collected with to describe trends in material behavior that would be desirable to depict 3 dimensions (Pilchak et al., 2013; 2016; Bridier et al., 2008; Clark et al., 2012; Germain et al., 2008).

The appeal of synthetic microstructures for tasks like mechanical/electrical property modeling variability prediction, and rare event characterization (Dawson & Boyce, 2015; Lebensohn et al., 2012; Tu et al., 2019), has motivated the creation of statistically-based microstructure generators (Quey et al., 2011; Groeber & Jackson, 2014), but these methods present challenges in terms of generation time and imitation of local statistical features (Hémery et al., 2019). As a result, there is growing interest in deep learning for synthetic microstructure generation (Robertson et al., 2023; Robertson & Kalidindi, 2022; Jangid et al., 2022b; Senthilnathan et al., 2021; Hsu et al., 2020; Brust et al., 2021; Kench & Cooper, 2021) using generative models, which can be used to generate complex material structures with lower cost burden. This materials problem has similarities in data structure/modality to other 2D and 3D computer vision tasks like robotics, LIDAR detection, and 3D image reconstruction, all of which share the challenge of having ground truth data that is difficult to collect. Significant advancements in deep generative models have been made in computer vision (Croitoru et al., 2023; Rombach et al., 2022), but a fundamental requirement for designing deep generative models is the availability of large numbers of datasets for training the network. Unfortunately, collecting microstructural information is more expensive and time-consuming than conventional photography (Jangid et al., 2023; Charpagne et al., 2021; Stinville et al., 2022; Rowen-

*Equal contribution [1]Electrical and Computer Engineering, University of California Santa Barbara [2]Mechanical Engineering, University of California Santa Barbara [3]Materials Department, University of California Santa Barbara. Correspondence to: Devendra Jangid <dkjangid@ucsb.edu>.

Accepted after peer-review at the 1st workshop on Synergy of Scientific and Machine Learning Modeling, SynS & ML ICML, Honolulu, Hawaii, USA. July, 2023. Copyright 2023 by the author(s).

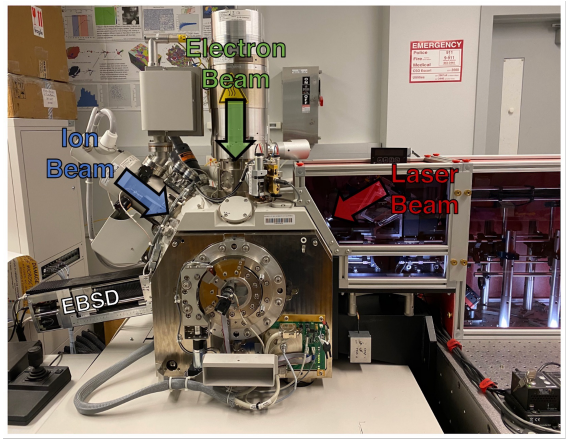

Figure 1. TRIBEAM MICROSCOPE: 3D microstructure datasets are obtained during a serial sectioning experiment. The TriBeam is so named due to the presence of three beams: the electron, focused ion, and laser. The electron beam is used to gather diffraction patterns, whereas the laser beam is used to ablate the materials, layer-by-layer. The ion beam can be used clean the laser ablated surface in damage sensitive materials, if necessary.

horst et al., 2020; DeCost et al., 2017), as the most basic light optical techniques still require careful material surface preparation (e.g. polishing, chemical etching) for just a single 2D image. Furthermore, these light optical techniques cannot capture richer microstructure information like crystallographic orientations and local chemistry, which require even costlier and more involved methods.

Electron Back Scatter Diffraction (EBSD) (Schwartz et al., 2009) is a widely used scanning electron microscopy (SEM) technique that images a material surface to collect microstructure data that contains crystal orientation information. Microstructure data describes the arrangement of crystalline structures in a material, where "crystalline" refers to a collection of atoms exhibiting long-range periodic order. For crystalline materials, the arrangement of atoms significantly influences several material properties including melting temperature, yield strength, ductility, and fatigue resistance. Light optical microscopy and the resulting RGB images have inadequate resolution to image atoms or to evaluate atomic arrangement. Therefore, we rely on X-ray or electron diffraction to understand the arrangement. During EBSD, a material is imaged one pixel at a time using a SEM, and electrons are diffracted from the atoms in the crystal according to Bragg's law. These diffracted electrons are collected on a detector producing a pattern called Kikuchi bands, which are indexed into crystal orientations to extract information about crystallographic arrangement. There are several techniques to determine crystal orientation from a Kikuchi pattern, including mathematical approximations of band locations using Hough or Radon transforms (Krieger Lassen et al., 1992; Adams et al., 1993), dictionary-based spherical cross-correlation approaches that leverage simulations of the electron-materials interactions (Jackson et al., 2019; Lenthe et al., 2019), and network-based approaches (Ding et al., 2021). The resultant data during EBSD is a diffraction pattern at each pixel in an image, which is indexed (mapped) into a crystal orientation represented as a vector (e.g. quaternion, Euler angle).

One of the greatest limitations of both light optical microscopy and EBSD is that they can only gather 2D images of materials that are fundamentally 3D. For both material and component design, 3D microstructure data plays a critical role in property prediction by informing the connectivity of grains and crystalline phases and characteristics of their interfaces for a broad range of applications, from biomedical to aerospace (National Research Council, 2008). As a result, 3D microscopy techniques, like the TriBeam depicted in figure 1, have been developed for this purpose. The TriBeam microscope is used to remove material in a layer-by-layer fashion from the sample using a femtosecond pulse laser, capturing images at the surface of each slice using an electron beam and a suite of detectors (Echlin et al., 2021). These slices are then aligned, indexed, and assembled into a 3D volume following the pipeline process shown in figure 2. The TriBeam method has made 3D microstructure collection more accessible (Echlin et al., 2021; Randolph et al., 2018), but even with enhancements in detector speeds and laser material removal rates, gathering this information remains expensive, energy-intensive, and time-consuming.

Here we release for the machine learning community an experimental microstructural dataset of a titanium alloy, Ti-6Al-4V on BisQue (`https://bit.ly/Ti6Al4Vmicrostructure`), a web-based cloud platform. We discuss the representation methods and physics needed to describe this microstructural dataset and design physics-based generative models for creating new synthetic microstructures. Our motivation for introducing 3D microstructure data to the machine learning community is to empower the development of different physics-based deep generative models for synthesizing microstructures, which can be used to design new materials.

2. Microstructure Dataset

The Ti-6Al-4V material released here (`https://bit.ly/Ti6Al4Vmicrostructure`) was obtained as a standard grade, double-melt plate, with a thickness of approximately 0.5 inches (1.25 cm). TriBeam characterization was used to construct 3D orientation maps of the heat-treated equiaxed microstructure, as shown in figure 3. The dataset resolution in the EBSD imaging plane is 0.6 micrometers (x and y axis), and in the sectioning plane is 1 micrometer (z axis). Polycrystalline materials like Ti-6Al-4V consist of a large number of subdomains called grains,

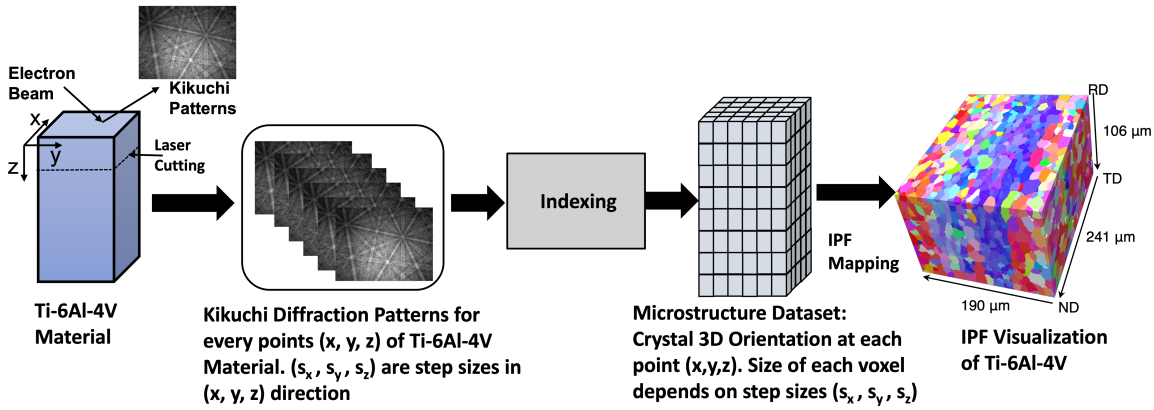

Figure 2. MICROSTRUCTURE DATA COLLECTION PROCESS: Material researchers collect Kikuchi diffraction patterns for each (x, y, z) coordinate of materials with the given step sizes (s_x, s_y, s_z) in (x, y, z) axis. The Kikuchi diffraction patterns are then indexed to determine the crystal orientation at each point. This information about crystal orientation is useful for predicting material properties. The inverse Pole Figure (IPF) technique is used to visualize the 3D crystal orientation.

which are visible as regions of relatively uniform crystallographic orientation (uniform color) in figure 3. The dataset presented here contains a total of 8893 grains, with 6645 unbiased grains (grains not at the volume edge). Differences in material composition and processing can change grain size (nm to mm range), morphology, and the distribution of orientations in grain neighborhoods. The arrangement and morphology of grains (and their orientations) can be correlated with material properties like strength and conductivity. For comparison, other 3D datasets of different material systems exist in the literature that are also openly accessible for use by this community (Stinville et al., 2022; Shade et al., 2019).

In the remainder of this section, we will discuss the fundamentals of microstructure and orientation representation from a computational perspective.

2.1. Orientation Representation

The orientation of a 3D point can be represented in many ways, including Euler angles, rotation matrices, quaternions, axis-angle pairs, and Rodrigues vectors. Each orientation representation has distinct advantages and disadvantages in terms of ease of use for different calculations and data visualization, and there are packages available to readily convert between them (Rowenhorst et al., 2015; De Graef). Within these possible representations, quaternions are frequently used to avoid ambiguity in 3D rotations (gimbal lock) in tasks like software graphics, computer vision, and robotics.

Quaternions: A quaternion q is a four component number of the form $q = q_0 + iq_1 + jq_2 + kq_3$, where imaginary units (i, j, k) satisfy the following relationship:

$$i^2 = j^2 = k^2 = ijk - 1 \tag{1}$$

Unit quaternions can always be written in the form

$$q = \cos\frac{\omega}{2} + \sin\frac{\omega}{2}(c_1 i + c_2 j + c_3 k) \tag{2}$$

where c_i are the directions cosines of the rotation axis unit vector $\hat{n}$. Unit quaternions are located on the sphere S_3 inside the 4D quaternion space.

Quaternion representation is preferable due to its simplicity in computing orientation differences (misorienation) and the ease of enforcing constraints for a valid unit quaternion rotation. Furthermore, the only redundancy in quaternion is that $q = -q$, which is computationally trivial to address. The efficiency and lack of ambiguity in quaternion representations also make them well-suited to orientation expression in neural networks, both for loss functions (Jangid et al., 2022a) and network layer design (Parcollet et al., 2018).

2.2. Symmetry Space

Optically-gathered images commonly used in computer vision tasks encode color, intensity, or topographic information as scalars at each pixel. On the other hand, crystallographic data encodes orientations as vector-based rotations relative to a chosen reference frame. The range of possible unique crystal rotations is constrained by the symmetry of the crystal, such as body centered cubic (BCC), face centered cubic (FCC), or hexagonal close packed (HCP). Symmetry has a disruptive impact on learning with standard loss functions (L-norms) because symmetry planes create duplicates and discontinuities in rotation distance measurement. Multiple-defined rotations are avoided through the fundamental zone convention, which is defined as the polyhedron in Rodrigues space that encompasses all angles whose distance is closer to the 0 rotation than to any symmetric equivalent of the 0 rotation (Morawiec & Field, 1996). The

Ti-6Al-4V microstructure dataset released here is an HCP material (space group 194, point group $6/mmm$), so it has a total of 24 symmetry operators, of which only 12 do not involve a change of handedness.

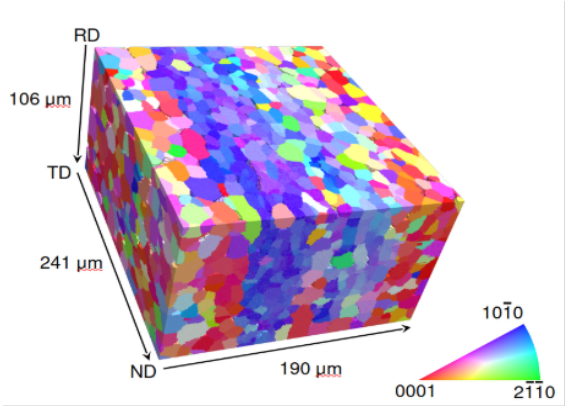

Figure 3. INVERSE POLE FIGURE (IPF) MAP OF TI-6AL-4V-EQUIAXED MICROSTRUCTURE DATASET: The color represents the orientation of crystal. Each grain of the same color has a nearly identical crystallographic orientation. Rolling direction (RD), transverse direction (TD) and normal direction (ND) reference arrows are also shown.

2.3. Visualization of Crystal Orientations

When orientation data is produced, real or synthetic, a means of visualization is required, which is challenging, since many orientation representations (quaternion, rotation matrix, Rodrigues vector) do not directly translate into common image formats. The most direct option is to convert orientation output into a 3-channel form and map it directly to a color scheme (RGB, HSV, YUV), but the presence of symmetry means small orientation changes often result in large changes in color scale in this type of mapping, which can create the appearance of noise even when data is correct. For this reason, the Inverse Pole Figure (IPF) color scheme was designed, which stereographically projects the fundamental zone into 2D and maps a uniform RGB gradient onto it. An IPF legend for HCP is shown in the bottom right of figure 3, and details of projection and color pattern choice are discussed in (Schwartz et al., 2009; Jackson et al., 2019; Nolze & Hielscher, 2016). Even though this output is intuitive to visualize, it is a projection that is fundamentally ambiguous and not information preserving, so no learning or inference should be done on IPF images as they cannot be converted back into orientations.

3. Incorporation of Physics of Microstructure Data into Generative Models:

The 3D arrangement of grains, phases, and interfaces controls material properties, but collecting this information experimentally remains costly. Despite this, one advantage of microstructure data over other 3D data is that its appearance is inherently dictated by physical relationships that can be incorporated into generative models to enable more realistic output and reduce data burdens. Models like Generative Adversarial Networks (GANs) have already been designed to explore some aspects of 3D microstructure, including grain and phase distribution, to varying degrees of success (Hsu et al., 2020; Kench & Cooper, 2021; Jangid et al., 2022b), but many open problems remain. Here, we present some examples of physics-based relationships that can be applied to generative models for synthetic microstructures, including as GANs, variational autoencoders, and diffusion models.

Crystal Symmetry: As previously discussed, symmetry space constrains distance metrics, but can also be used to simplify orientation representation space. One approach is to incorporate symmetry directly into the loss function, which is a softer constraint that does not increase inference time (Jangid et al., 2022a). In a symmetry-based loss, the orientation at every pixel is considered as a collection of equivalent rotations across all symmetries and the loss distance is calculated as the minimum distance between the ground truth value and any value within this collection. Another approach is to introduce symmetry directly into the convolution kernels (Cohen et al., 2018; 2019). This approach is a stronger constraint, but may also increase the inference time for the network.

Rotational Distance: The microstructure data contains information about the 3D orientation of crystals as normalized vector expressions of rotations (unit quaternion rotations are located on the $\mathbb{S}^3$ manifold in $\mathbb{R}^4$). Regardless of expression form, when evaluating network output against ground truth, the most physically accurate distance metric is the geodesic distance along the rotation sphere, rather than L-norms. The rotational distance between two quaternions can be computed as the following:

$$\theta = 4\sin^{-1}\left(\frac{d_{\text{euclid}}}{2}\right) \tag{3}$$

where, $d_{\text{euclid}} = \|q_1 - q_2\|_2$. This approach requires some approximations to resolve the discontinuities in the inverse sine function, but it is appealing because it allows for direct comparison of network error and experimental results.

Incorporation of Data Modalities into Network Architecture: Similar to symmetry, crystal orientation representation can be incorporated directly into network architectures to improve performance and reduce complexity/data burdens. For quaternions, approaches to this include the quaternion convolution kernel (Parcollet et al., 2018) and quaternion transformer (Jangid et al., 2023). As an example, quaternion convolution computes the Hamilton product between input feature maps and kernel filters to better conserve data representation. However, introducing non-linear activations

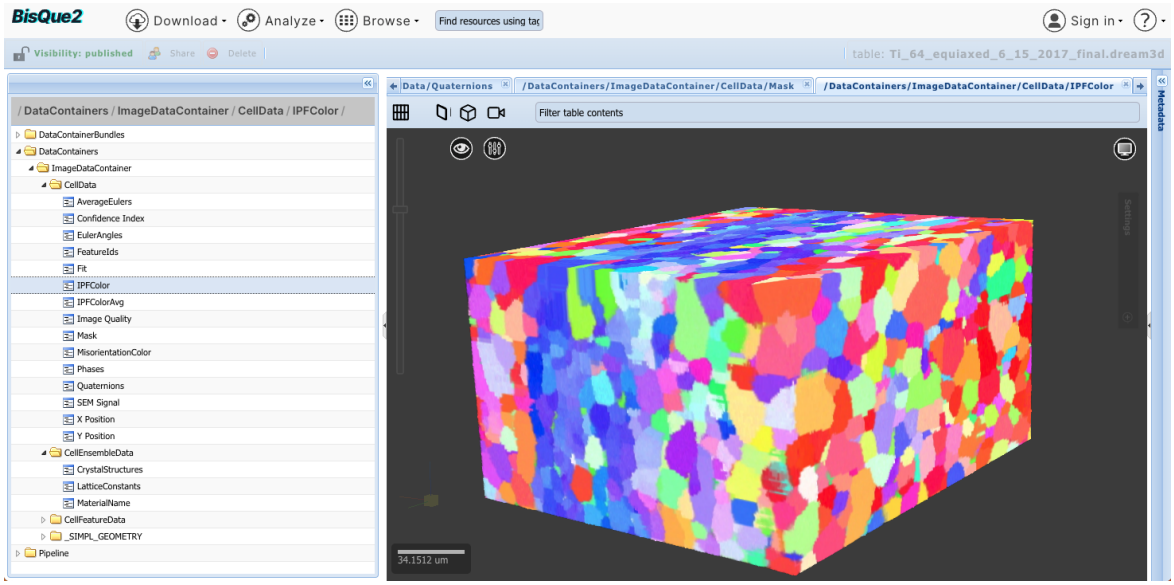

Figure 4. TI-6AL-4V MICROSTRUCTURE DATASET ON BISQUE: Ti-6Al-4V 3D microstructure dataset can be accessed on our open source web-based architecture "BisQue". The microstructure dataset are saved as hdf/dream3d format and can visualized using "IPFColor" field. The crystal orientations at each voxel is saved in "EulerAngles" and "Quaternions" fields. "FeatureIds" field shows the location of grains, and each grain has a distinct feature id.

is not straightforward, as the only functions that satisfy the Cauchy-Riemann-Fueter equations for quaternions are linear or constant (Parcollet et al., 2020). Locally analytic quaternion activation functions compatible with standard backpropagation have been developed (De Leo & Rotelli, 1997; Isokawa et al., 2012), but the most straightforward way to introduce non-linearity is element-wise, which does not preserve inter-channel relationship.

4. Accessing 3D Microstructure Dataset:

In this section, the details of Ti-6Al-4V microstructure dataset, available on "BisQue" (`https://bit.ly/Ti6Al4Vmicrostructure`), are discussed. The physical dimension of the Ti-6Al-4V microstructure sample is $190 \times 241 \times 106$ micrometers. In voxel-based representation, this 3d volume expands to $318 \times 377 \times 121$ voxels, where each voxel contains information about the crystal orientation. All the relevant information for designing machine learning models relevant to this dataset is stored in DataContainers/ImageDataContainer/CellData, as shown in the figure 4. Each grain in the dataset is composed of a collection of neighboring voxels that have similar crystal orientations. In the dataset itself, each grain is labeled with a unique positive integer value, termed a feature ID, that is randomly assigned, starting at 1 (the integer value 0 is typically reserved for any void regions captured in the dataset). There are a total of 8893 single grains in the $318 \times 377 \times 121$ volume, and of those, 6645 grains are completely captured such that they do not come into contact with the edge of the collection volume. Separating fully captured grains from edge grains is not usually necessary for voxel-based vision tasks, but is important for morphological studies where flat surfaces created by sample edges can lead to bias. The exact number of datapoints in the dataset will vary depending on the features of interest being studied. Any voxel level information will be stored in an array of $318 \times 377 \times 121 \times$ [array depth]. Quaternion information, for example, is collected at the voxel level, so the array of quaternions in this dataset is of size $318 \times 377 \times 121 \times 4$. However, there is also information that can be collected at the grain level (also referred to as feature level), such as the average orientation in each grain, or the volume of each grain. Values like these would be stored as a list, with the value for a given grain stored at the list index that matches its integer feature ID.

In orientation space, the problem of generating synthetic microstructures can be subdivided into the two distinct challenges of 2D and 3D microstructure generation. While 3D microstructure generation holds greater value for researchers, designing physics-based generative models for 3D is hindered by the limited availability of real 3D data. As a preliminary step towards working with 3D microstructures, 2D microstructure generation is often an essential task. Because 3D EBSD datasets do not have any visual perspective associated with them, they can readily be broken up into a collection of correlated slices along orthogonal directions for the purposes of 2D learning.

Broader Impact

Since the 1990s, there have been several large-scale manufacturing initiatives, including the Materials Genome Initiative (MGI) (National Science and Technology Council (US), 2011) and Integrated Computational Materials Engineering (National Research Council, 2008), whose goal has been accelerating the materials development process. A key bottleneck of the current design process is the slow pace of experimentation, motivating for simulations that can replace or better target these costly steps. Generative models for 3D microstructures can accelerate this process by providing an efficient means of producing realistic input to engineering component-scale models. With good generative models, a single 3D microstructure dataset can be used to create and test many synthetic variations, effectively replacing hundreds of costly experiments. Generative models for 3D microstructure also broaden research access to 3D microstructures, which are difficult to gather experimentally. Finally, physics-based network architectures built for crystallographic rotations may also have translational applicability in other fields that are dependent on rotational inference.

Acknowledgements

This research is supported in part by NSF award number 1664172. N. Brodnik gratefully acknowledges financial support from NSWC Grant (N00174-22-1-0020). The authors gratefully acknowledge Patrick Callahan, Toby Francis, Andrew Polonsky, and Joseph Wendorf for aiding in the collection of the 3D Ti-6Al-4V datasets. The MRL Shared Experimental Facilities are supported by the MRSEC Program of the NSF under Award No. DMR 1720256; a member of the NSF-funded Materials Research Facilities Network (www.mrfn.org). Use was also made of computational facilities purchased with funds from the National Science Foundation (CNS-1725797) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara. Use was made of the computational facilities purchased with funds from the National Science Foundation CC* Compute grant (OAC-1925717) and administered by the Center for Scientific Computing (CSC). The ONR Grant N00014-19-2129 is also acknowledged for the titanium datasets. The authors of this work declare no competing financial or non-financial interests.

References

Adams, B. L., Wright, S. I., and Kunze, K. Orientation imaging: The emergence of a new microscopy. *Metallurgical Transactions A*, 24(4):819–831, April 1993. doi: 10.1007/bf02656503. URL https://doi.org/10.1007/bf02656503.

Bridier, F., Villechaise, P., and Mendez, J. Slip and fatigue crack formation processes in an / titanium alloy in relation to crystallographic texture on different scales. *Acta Materialia*, 56(15):3951–3962, 2008. ISSN 1359-6454. doi: https://doi.org/10.1016/j.actamat.2008.04.036. URL https://www.sciencedirect.com/science/article/pii/S1359645408002917.

Brust, A., Payton, E., Hobbs, T., Sinha, V., Yardley, V., and Niezgoda, S. Probabilistic reconstruction of austenite microstructure from electron backscatter diffraction observations of martensite. *Microscopy and Microanalysis*, 27(5):1035–1055, September 2021. doi: 10.1017/s1431927621012484. URL https://doi.org/10.1017/s1431927621012484.

Charpagne, M., Stinville, J. C., Polonsky, A. T., Echlin, M. P., and Pollock, T. M. A multi-modal data merging framework for correlative investigation of strain localization in three dimensions. *JOM*, 73(11):3263–3271, September 2021. doi: 10.1007/s11837-021-04894-6. URL https://doi.org/10.1007/s11837-021-04894-6.

Clark, D., Whittaker, M. T., and Bache, M. R. Microstructural characterization of a prototype titanium alloy structure processed via direct laser deposition (dld). *Metallurgical and Materials Transactions B*, 43:388–396, 2012.

Cohen, T., Weiler, M., Kicanaoglu, B., and Welling, M. Gauge equivariant convolutional networks and the icosahedral CNN. In Chaudhuri, K. and Salakhutdinov, R. (eds.), *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 1321–1330. PMLR, 09–15 Jun 2019. URL https://proceedings.mlr.press/v97/cohen19d.html.

Cohen, T. S., Geiger, M., Koehler, J., and Welling, M. Spherical cnns, 2018.

Croitoru, F.-A., Hondru, V., Ionescu, R. T., and Shah, M. Diffusion models in vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

Dawson, P. R. and Boyce, D. E. FEpX – Finite Element Polycrystals: Theory, Finite Element Formulation, Numerical Implementation and Illustrative Examples, April 2015. URL http://arxiv.org/abs/1504.03296. arXiv:1504.03296 [cond-mat].

De Graef, M. Marc degraef's 3d rotations github repository. URL https://github.com/marcdegraef/3Drotations. Accessed: 2023-05.

De Leo, S. and Rotelli, P. Local hypercomplex analyticity. *arXiv preprint funct-an/9703002*, 1997.

DeCost, B. L., Francis, T., and Holm, E. A. Exploring the microstructure manifold: image texture representations applied to ultrahigh carbon steel microstructures. *Acta Materialia*, 133:30–40, 2017.

DeMott, R., Collins, P., Kong, C., Liao, X., Ringer, S., and Primig, S. 3d electron backscatter diffraction study of α lath morphology in additively manufactured ti-6al-4v. *Ultramicroscopy*, 218:113073, 2020.

DeMott, R., Haghdadi, N., Kong, C., Gandomkar, Z., Kenney, M., Collins, P., and Primig, S. 3d electron backscatter diffraction characterization of fine titanium microstructures: collection, reconstruction, and analysis methods. *Ultramicroscopy*, 230:113394, 2021. ISSN 0304-3991. doi: https://doi.org/10.1016/j.ultramic.2021.113394. URL https://www.sciencedirect.com/science/article/pii/S0304399121001741.

Ding, Z., Zhu, C., and Graef, M. D. Determining crystallographic orientation via hybrid convolutional neural network. *Materials Characterization*, 178:111213, August 2021. doi: 10.1016/j.matchar.2021.111213. URL https://doi.org/10.1016/j.matchar.2021.111213.

Echlin, M. P., Polonsky, A. T., Lamb, J., Geurts, R., Randolph, S. J., Botman, A., and Pollock, T. M. Recent developments in femtosecond laser-enabled TriBeam systems. *JOM*, 73(12):4258–4269, November 2021. doi: 10.1007/s11837-021-04919-0. URL https://doi.org/10.1007/s11837-021-04919-0.

Germain, L., Gey, N., Humbert, M., Vo, P., Jahazi, M., and Bocher, P. Texture heterogeneities induced by subtransus processing of near α titanium alloys. *Acta Materialia*, 56 (16):4298–4308, 2008.

Groeber, M. A. and Jackson, M. A. Dream. 3d: a digital representation environment for the analysis of microstructure in 3d. *Integrating materials and manufacturing innovation*, 3(1):56–72, 2014.

Hémery, S., Naït-Ali, A., Guéguen, M., Wendorf, J., Polonsky, A., Echlin, M., Stinville, J., Pollock, T., and Villechaise, P. A 3d analysis of the onset of slip activity in relation to the degree of micro-texture in ti–6al–4v. *Acta Materialia*, 181:36–48, December 2019. doi: 10.1016/j.actamat.2019.09.028. URL https://doi.org/10.1016/j.actamat.2019.09.028.

Hsu, T., Epting, W. K., Kim, H., Abernathy, H. W., Hackett, G. A., Rollett, A. D., Salvador, P. A., and Holm, E. A. Microstructure generation via generative adversarial network for heterogeneous, topologically complex 3d materials. *JOM*, 73(1):90–102, December 2020. doi: 10.1007/s11837-020-04484-y. URL https://doi.org/10.1007/s11837-020-04484-y.

Isokawa, T., Nishimura, H., and Matsui, N. Quaternionic multilayer perceptron with local analyticity. *Information*, 3(4):756–770, 2012.

Jackson, M. A., Pascal, E., and Graef, M. D. Dictionary indexing of electron back-scatter diffraction patterns: a hands-on tutorial. *Integrating Materials and Manufacturing Innovation*, 8(2):226–246, May 2019. doi: 10.1007/s40192-019-00137-4. URL https://doi.org/10.1007/s40192-019-00137-4.

Jangid, D. K., Brodnik, N. R., Goebel, M. G., Khan, A., Majeti, S., Echlin, M. P., Daly, S. H., Pollock, T. M., and Manjunath, B. Adaptable physics-based super-resolution for electron backscatter diffraction maps. *npj Computational Materials*, 8(1):255, 2022a.

Jangid, D. K., Brodnik, N. R., Khan, A., Goebel, M. G., Echlin, M. P., Pollock, T. M., Daly, S. H., and Manjunath, B. 3d grain shape generation in polycrystals using generative adversarial networks. *Integrating Materials and Manufacturing Innovation*, 11:71–84, 2022b.

Jangid, D. K., Brodnik, N. R., Echlin, M. P., Pollock, T. M., Daly, S. H., and Manjunath, B. Q-rbsa: High-resolution 3d ebsd map generation using an efficient quaternion transformer network. *arXiv preprint arXiv:2303.10722*, 2023.

Kelly, M. N., Glowinski, K., Nuhfer, N. T., and Rohrer, G. S. The five parameter grain boundary character distribution of -ti determined from three-dimensional orientation data. *Acta Materialia*, 111:22–30, 2016. ISSN 1359-6454. doi: https://doi.org/10.1016/j.actamat.2016.03.029. URL https://www.sciencedirect.com/science/article/pii/S1359645416301811.

Kench, S. and Cooper, S. J. Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion. *Nature Machine Intelligence*, 3(4):299–305, April 2021. doi: 10.1038/s42256-021-00322-1. URL https://doi.org/10.1038/s42256-021-00322-1.

Krieger Lassen, N. C., Juul Jensen, D., and Conradsen, K. Image processing procedures for analysis of electron back scattering patterns. *Scanning microscopy*, 6(1):115–121, 1992.

Lebensohn, R. A., Kanjarla, A. K., and Eisenlohr, P. An elasto-viscoplastic formulation based on fast Fourier transforms for the prediction of micromechanical fields in polycrystalline materials. *International Journal of Plasticity*, 32-33:59–69, May

2012. ISSN 07496419. doi: 10.1016/j.ijplas.2011.12.005. URL https://linkinghub.elsevier.com/retrieve/pii/S0749641911001951.

Lenthe, W., Singh, S., and Graef, M. D. A spherical harmonic transform approach to the indexing of electron back-scattered diffraction patterns. *Ultramicroscopy*, 207:112841, December 2019. doi: 10.1016/j.ultramic.2019.112841. URL https://doi.org/10.1016/j.ultramic.2019.112841.

Morawiec, A. and Field, D. Rodrigues parameterization for orientation and misorientation distributions. *Philosophical Magazine A*, 73(4):1113–1130, 1996.

National Research Council. *Integrated Computational Materials Engineering: A Transformational Discipline for Improved Competitiveness and National Security*. The National Academies Press, Washington, DC, 2008. ISBN 978-0-309-11999-3. doi: 10.17226/12199.

National Science and Technology Council (US). *Materials genome initiative for global competitiveness*. Executive Office of the President, National Science and Technology Council, 2011.

Nolze, G. and Hielscher, R. Orientations – perfectly colored. *Journal of Applied Crystallography*, 49 (5):1786–1802, September 2016. doi: 10.1107/s1600576716012942. URL https://doi.org/10.1107/s1600576716012942.

Parcollet, T., Zhang, Y., Morchid, M., Trabelsi, C., Linarès, G., De Mori, R., and Bengio, Y. Quaternion convolutional neural networks for end-to-end automatic speech recognition. *arXiv preprint arXiv:1806.07789*, 2018.

Parcollet, T., Morchid, M., and Linarès, G. A survey of quaternion neural networks. *Artificial Intelligence Review*, 53(4):2957–2982, 2020.

Patala, S., Mason, J. K., and Schuh, C. A. Improved representations of misorientation information for grain boundary science and engineering. *Progress in Materials Science*, 57:1383–1425, 2012. doi: 10.1016/j.pmatsci.2012.04.002. URL http://dx.doi.org/10.1016/j.pmatsci.2012.04.002.

Pilchak, A., Szczepanski, C., Shaffer, J., Salem, A., and Semiatin, S. Characterization of microstructure, texture, and microtexture in near-alpha titanium mill products. *Metallurgical and Materials Transactions A*, 44:4881–4890, 2013.

Pilchak, A. L., Shank, J., Tucker, J. C., Srivatsa, S., Fagin, P. N., and Semiatin, S. L. A dataset for the development, verification, and validation of microstructure-sensitive process models for near-alpha titanium alloys. *Integrating Materials and Manufacturing Innovation*, 5(1):259–276, Dec 2016. ISSN 2193-9772. doi: 10.1186/s40192-016-0056-1.

Quey, R., Dawson, P., and Barbe, F. Large-scale 3d random polycrystals for the finite element method: Generation, meshing and remeshing. *Computer Methods in Applied Mechanics and Engineering*, 200(17-20):1729–1745, 2011.

Randolph, S. J., Filevich, J., Botman, A., Gannon, R., Rue, C., and Straw, M. In situ femtosecond pulse laser ablation for large volume 3d analysis in scanning electron microscope systems. *Journal of Vacuum Science & Technology B*, 36(6):06JB01, November 2018. doi: 10.1116/1.5047806. URL https://doi.org/10.1116/1.5047806.

Robertson, A. E. and Kalidindi, S. R. Efficient generation of anisotropic n-field microstructures from 2-point statistics using multi-output gaussian random fields. *Acta Materialia*, 232:117927, 2022.

Robertson, A. E., Kelly, C., Buzzy, M., and Kalidindi, S. R. Local-global decompositions for conditional microstructure generation. *Acta Materialia*, pp. 118966, 2023.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10684–10695, 2022.

Rowenhorst, D., Lewis, A., and Spanos, G. Three-dimensional analysis of grain topology and interface curvature in a β-titanium alloy. *Acta Materialia*, 58(16): 5511–5519, 2010.

Rowenhorst, D., Rollett, A. D., Rohrer, G. S., Groeber, M., Jackson, M., Konijnenberg, P. J., and Graef, M. D. Consistent representations of and conversions between 3d rotations. *Modelling and Simulation in Materials Science and Engineering*, 23(8):083501, October 2015. doi: 10.1088/0965-0393/23/8/083501. URL https://doi.org/10.1088/0965-0393/23/8/083501.

Rowenhorst, D. J., Nguyen, L., Murphy-Leonard, A. D., and Fonda, R. W. Characterization of microstructure in additively manufactured 316l using automated serial sectioning. *Current Opinion in Solid State and Materials Science*, 24(3):100819, June 2020. doi: 10.1016/j.cossms.2020.100819. URL https://doi.org/10.1016/j.cossms.2020.100819.

Schwartz, A. J., Kumar, M., Adams, B. L., and Field, D. P. Electron backscatter diffraction in materials science. *Electron Backscatter Diffraction in Materials Science*, pp. 1–403, 2009. doi: 10.1007/978-0-387-88136-2.

Senthilnathan, A., Acar, P., and Graef, M. D. Markov random field based microstructure reconstruction using the principal image moments. *Materials Characterization*, 178:111281, August 2021. doi: 10.1016/j.matchar.2021.111281. URL https://doi.org/10.1016/j.matchar.2021.111281.

Shade, P. A., Musinski, W. D., Shah, M. N., Uchic, M. D., Donegan, S. P., Chapman, M. G., Park, J.-S., Bernier, J. V., Kenesei, P., Menasche, D. B., Obstalecki, M., Schwalbach, E. J., Miller, J. D., Groeber, M. A., and Cox, M. E. Afrl am modeling challenge series: Challenge 4 data package, 2019. URL https://petreldata.net/mdf/detail/shade_afrl_am_package_v2.1.

Stinville, J. C., Hestroffer, J. M., Charpagne, M. A., Polonsky, A. T., Echlin, M. P., Torbet, C. J., Valle, V., Nygren, K. E., Miller, M. P., Klaas, O., Loghin, A., Beyerlein, I. J., and Pollock, T. M. Multi-modal dataset of a polycrystalline metallic material: 3d microstructure and deformation fields. *Scientific Data*, 9(1), August 2022. doi: 10.1038/s41597-022-01525-w. URL https://doi.org/10.1038/s41597-022-01525-w.

Tu, X., Shahba, A., Shen, J., and Ghosh, S. Microstructure and property based statistically equivalent rves for polycrystalline-polyphase aluminum alloys. *International Journal of Plasticity*, 115:268–292, 2019.

A. Appendix

The EBSD dataset presented in this work was processed using Dream3D ((Groeber & Jackson, 2014)), so the structure of the set follows the HDF5 architecture and nomeclature conventions used for Dream3D files. Files with the .dream3d extension are directly compatible with applications and packages used for HDF files (e.g. the HDF View software package or h5py in Python), so these can readily be used to extract individual variables from the file for independent manipulation. The file structure can also be visualized directly within the web-based BisQue infrastructure for convenience. A visualization of the file directory from BisQue can be see in Figure 5, and a breakdown of variable definitions is available in Table 1.

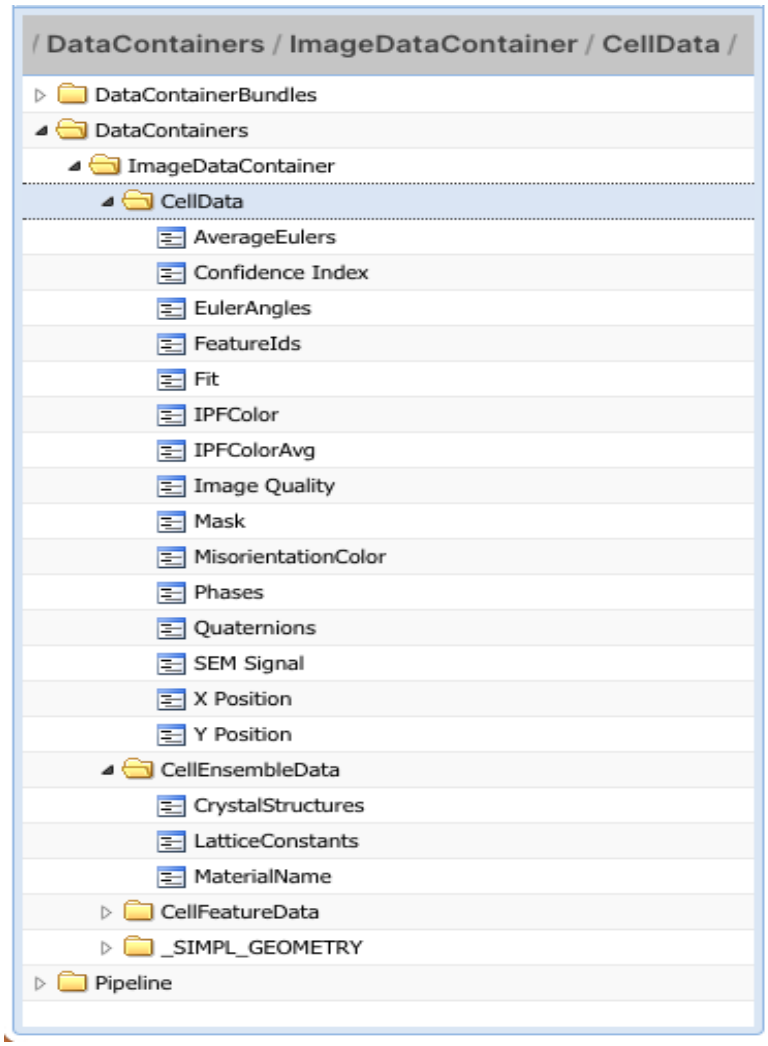

Figure 5. FILE STRUCTURE OF TI-6AL-4V MICROSTRUCTURE DATASET ON BISQUE: .

CellData	
AverageEulers	The average Euler angle for each voxel's corresponding grain (See AvgEulers).
Confidence Index	The confidence value for the Hough Transform indexing used to determine orientation.
Euler Angles	Crystal orientation at each voxel using Bunge Euler angle representation.
FeatureIDs	Integer identifier values for each grain. Void is feature 0. Others assigned randomly.
Fit	Comparative quality metric between the indexed orientation and Hough transform (in degrees).
IPFColor	RGB color mapping of crystal orientation using IPF projection
IPFColorAvg	IPF projection mapping where each grain is colored by average orientation.
Image Quality	Image quality metric. Larger values indicate better quality.
Mask	Mask indicating regions of void. Void = 0, Solid = 1 (note: dataset contains no void).
Misorientation Color	RGB color representation for relative misorientation, developed by (Patala et al., 2012).
Phases	Voxel level numerical phase map. 1 = HCP α phase. 2 = BCC β phase.
Quaternions	Crystal orientation at each voxel using quaternion representation.
SEM Signal	Dataframe for storing equivalent SEM signal. Dataset is purely EBSD, so this is empty.
X Position	The x-position of each voxel in the volume
Y Position	The y-position of each voxel in the volume.
CellEnsembleData	
CrystalStructures	Dream3D internal value. Numerical labels for void, α, and β crystal structures.
LatticeConstants	Lattice constants for associated labels in CrystalStructures
MaterialName	String name labels for phases in CrystalStructures and LatticeConstants.
CellFeatureData	
Active	Internal label indicating which features are captured within the current dataset.
AvgEulers	Average Euler angle orientation of each grain feature.
AvgQuats	Average quaternion orientation of each grain feature.
Centroids	Centroid of each grain in physical space (see SIMPL_Geometry).
EquivalentDiameters	Diameter in μm of each grain if it were approximated as a sphere of equivalent volume.
NumCells	Number of voxels in each grain feature.
NumNeighbors	Number of adjacent feature neighbors for each grain.
Volumes	The volume of each grain in μm^3
_SIMPL_Geometry	
Dimensions	Size of the dataset in voxels.
Origin	Location of the physical space origin in in μm.
Spacing	Physical x, y, and z size of each voxel in μm

Table 1. DESCRIPTION OF FILE STRUCTURE OF TI-6AL-4V MICROSTRUCTURE DATASET: