Convolutional Neural Networks for the Localization of Plastic Velocity Gradient Tensor in Polycrystalline Microstructures

David Montes de Oca Zapiain
George W. Woodruff School of Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332
e-mail: dmo7@gatech.edu.com

Apaar Shanker
School of Computational Science and Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332
e-mail: ashanker9@gatech.edu

Surya R. Kalidindi[1]
George W. Woodruff School of Mechanical Engineering,
School of Computational Science and Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332
e-mail: surya.kalidindi@me.gatech.edu

Recent work has demonstrated the potential of convolutional neural networks (CNNs) in producing low-computational cost surrogate models for the localization of mechanical fields in two-phase microstructures. The extension of the same CNNs to polycrystalline microstructures is hindered by the lack of an efficient formalism for the representation of the crystal lattice orientation in the input channels of the CNNs. In this paper, we demonstrate the benefits of using generalized spherical harmonics (GSH) for addressing this challenge. A CNN model was successfully trained to predict the local plastic velocity gradient fields in polycrystalline microstructures subjected to a macroscopically imposed loading condition. Specifically, it is demonstrated that the proposed approach improves significantly the accuracy of the CNN models when compared with the direct use of Bunge–Euler angles to represent the crystal orientations in the input channels. Since the proposed approach implicitly satisfies the expected crystal symmetries in the specification of the input microstructure to the CNN, it opens new research directions for the adoption of CNNs in addressing a broad range of polycrystalline microstructure design and optimization problems.
[DOI: 10.1115/1.4051085]

Introduction

Metals exhibit rich and complex granular microstructures, where the mechanical response of each grain is most naturally described in a crystal reference frame aligned with the motif describing the atomic configuration in the material. Such polycrystalline microstructures naturally produce heterogeneous mechanical (e.g., stress, strain) fields at the mesoscale when subjected to any macroscale loading, including those usually prescribed as uniform boundary conditions in multiscale simulations of the material response [1–4]. Efforts aimed at the design and development of advanced metals require low computational cost tools capable of accurately predicting the mesoscale mechanical fields controlling the performance characteristics of interest [5–8]. For example, the mesoscale stress and strain fields are known to influence strongly the fatigue and/or damage resilience of the metal [9,10]. One often needs to rapidly evaluate a large set of potentially viable microstructures before arriving at specific ones that exhibit optimal effective material properties or performance characteristics. In the context of polycrystalline microstructures, the design space of microstructures would include a very wide range of distributions of the crystal orientations (i.e., textures) as well as the size and shape distributions of the crystalline regions (i.e., grains). Therefore, there is a critical need for low-computational cost simulation tools capable of predicting accurately the mesoscale mechanical fields in polycrystalline microstructures of metals when subjected to macroscale loading conditions; such tools are essential for addressing microstructure design and optimization problems with typically available computational resources.

In established composite theories, the prediction of the mesoscale fields subjected to imposed macroscale loading conditions is generally referred to as the localization problem [11–14]. In essence, this is the inverse of the homogenization problem [8,11,15] which aims to compute the effective properties associated with a heterogeneous microstructure, which are typically needed for the design of engineered components. The solution to the localization problem needs to satisfy the governing equilibrium equations, the material constitutive laws at the mesoscale, and the imposed boundary conditions. These are essential in addressing material responses in fatigue and other failures [8,16,17]. In spite of the major strides made in the efforts aimed at developing analytical solutions to the localization problem [18–20], they are currently available only for a limited class of highly simplified or idealized material problems. For most practical problems of interest, the most versatile approach available for addressing the localization problems is to leverage numerical approaches based on the finite element (FE) method [21–24]. Although the FE-based strategies are known to produce highly accurate and reliable solutions to the localization problem, they often demand major computational resources. This is especially true when incorporating sophisticated physics related to the plastic response of the individual grains in the polycrystalline microstructure of the metals. As such, the FE-based approaches to the localization problems are not ideally suited for addressing microstructure design problems, where one expects to evaluate an exceedingly large number of potential microstructures in order to identify the optimized solution. The materials design challenges described above are further exacerbated by the recent advances in multiscale materials modeling approaches [25–27], which have now made it possible to consider the rich and more complex details of the material structure at multiple length scales in evaluating the material's response. However, in order to transform these advances into materials with improved properties, one needs novel, practically viable, computational strategies for exploring large materials design spaces spanning the acceptable ranges of suitably defined measures of the multiscale material internal structure.

It is vital that these new computational approaches provide solutions within the constraints of the computational resources typically available to the materials design and manufacturing industry.

The advent of artificial intelligence/machine learning approaches has opened new research avenues for addressing the gap identified above. Specifically, it has raised the prospect of establishing sufficiently accurate low-computational cost surrogate models trained to the results obtained from the computationally expensive FE-based models in a one-time training exercise using a collection of digitally created microstructures and subsequently using the surrogate models for predicting the mesoscale mechanical fields in new microstructures. Implementation of such a strategy requires suitable customization of the tools developed originally for other applications in computer science (e.g., image recognition [28,29]), which are quite unrelated to the concepts and paradigms employed in the mechanics of materials field. In other words, it should be possible to leverage some of the emergent concepts and tools from the field of machine learning to distill transferable knowledge from FE simulations conducted on a set of training microstructures and subsequently use the captured knowledge to predict the response of new (test) microstructures. Although the training phase would incur a one-time substantial computational cost, all subsequent predictions for new microstructures can be obtained with minimal computational cost. A specific example of the implementation of this paradigm can be seen in the recently formulated materials knowledge systems (MKS) framework. The initial efforts in the development of the MKS framework adopted the parametric model forms suggested by the statistical continuum theories of Kröner [19,20] and employed linear regression techniques to train the surrogate models [30–35]. This simple strategy was found to produce very accurate localization surrogate models for material systems with low to moderate contrast in the local properties of the constituents. As an example, in a two-phase microstructure with isotropic constituents, the ratio of the yield strengths of the constituent phases would reflect the contrast in the composite material system. Prior work in the MKS framework [11,33,35] has demonstrated that simple linear expansions can be used effectively to obtain highly accurate localization surrogate models for low-contrast composites. High-contrast composite material systems exhibit significantly higher levels of heterogeneity in the mesoscale mechanical fields (e.g., stress, strain) and demand more sophisticated approaches [8,34,36].

Modern machine learning tools have opened new avenues for building highly accurate localization surrogate models for high-contrast composite material systems. In recent work [37,38], it was especially noted that convolutional neural networks (CNNs) offer a powerful tool set to address this challenge because the use of convolutions in the linear transformations employed in the CNN is indeed very similar to the convolution operators found in the analytical solutions derived for the localization problems. Furthermore, the localization solutions provided by CNNs implicitly exhibit translational invariance, which is a key property expected in all localization solutions based on the governing physics for this class of problems. CNNs employ layers of convolutional filters along with nonlinear activations to efficiently and accurately capture the complex neighborhood interactions that occur naturally at the mesoscale, when the heterogeneous material microstructure is subjected to a prescribed macroscale loading condition. The convolution filter weights in the CNN are inferred by training it to ground-truth data, which can be obtained in a one-time effort by executing suitable FE simulations on an ensemble of digitally created microstructures [39]. CNNs have already been successfully employed in obtaining highly accurate localization solutions for elastic deformations in high-contrast two-phase composites [38].

As the logical next step in the further development of CNN-based localization surrogate models, there is a critical need to explore their viability for more complex microstructures and more sophisticated governing physics. Specifically, we explore in this paper the viability of CNN-based localization models for plastic deformations in polycrystalline metal samples subjected to a prescribed macroscale imposed loading condition. The governing physics for these problems is specified by the crystal plasticity theories [1–4,21,40,41] which are significantly more complex than the elastic deformations studied in the prior applications of CNNs to localization problems. Another major challenge comes from the need to describe the local material state in each voxel of the polycrystalline microstructure, which is most commonly expressed as a set of rotation angles known as the Bunge–Euler angles, $g = (\varphi_1, \Phi, \varphi_2)$ [42]. This set of angles represents the local crystal lattice orientation in each voxel of the representative polycrystalline sample volume with respect to the sample reference frame. Mathematically, g is an element of SO(3) space, and its representation in the input channels to the CNN poses significant challenges. Because of the inherent crystal symmetries, any specified crystal orientation is expected to have multiple equivalent representations. For example, each orientation of a cubic crystal would have 24 equivalent descriptors in SO(3). Furthermore, orientation spaces defined using representations based on Bunge–Euler angles exhibit multiple degeneracies; as a result, the Euclidean distances in such spaces do not reflect accurately the similarity or closeness of any two orientations selected in SO(3). Established crystal plasticity theories [1–4,21,40,41] used in the physics-based approaches are formulated to implicitly and accurately reflect the necessary attributes of the representation of the orientations by formulating and employing appropriate coordinate transformation matrices and the expression of physical laws in ways that automatically reflect the expected crystal symmetries. Consequently, it is necessary to develop suitable approaches for incorporating these considerations in the formulation of CNNs for the localization surrogate models applicable to polycrystalline representative volume elements (RVEs). Of course, one may allow the CNNs to learn the equivalencies, periodicity, and the distortions inherent to the SO(3) space by themselves. However, such an approach is likely to be highly inefficient and demand a significantly higher training cost. Current applications of deep learning-based surrogate models for polycrystalline materials simply treat the differently oriented grains as different phases of a multiphase composite [43–45] and do not address the challenges identified above.

This work addresses the central challenge described above by transforming the crystal lattice orientation in each voxel of the three-dimensional (3D) polycrystalline microstructure into the Fourier coefficients of a suitably defined orientation distribution function. This is accomplished using the generalized spherical harmonics (GSH) [42] as the Fourier basis. GSH automatically reflect the desired crystal symmetries and provide a much more meaningful similarity measures for the elements of the SO(3) space. GSH have indeed been successfully employed in related problems [30,31,34,46], where they have been shown to produce tremendous economy in representing the salient features of the local material state. In this work, we explore the potential benefits of using the same GSH representations of the polycrystalline microstructure for developing CNN surrogate models. More specifically, it will be shown that the novel GSH-CNNs can be trained to produce highly accurate localization surrogate models for predicting the components of the local plastic velocity gradient tensor in the individual voxels of 3D polycrystalline RVEs. It is also demonstrated that the accuracy of the GSH-CNN significantly outperforms a CNN trained on the input microstructures described directly using the Bunge–Euler angles, while also significantly outperforming the first-order MKS surrogate model that serves as the current standard [34].

Background

Microstructure Quantification. In order to quantify the internal structure of polycrystalline materials, we first discretize the spatial domain of the RVE with a uniform grid of spatial cells (i.e., voxels), enumerated by $s \in S$ (typically used as a vector index for three-dimensional RVEs [11]). We then discretize the potential material local states that can occupy the spatial voxel s in a similar manner with $h \in H$. As a result, we can define a discretized microstructure function, denoted conveniently by the array

m_s^h, whose components express the volume fraction of the material local state h present in voxel s [8,11]. This simple notion of m_s^h is directly applicable for multiphase microstructures with two or more thermodynamically distinct phases (enumerated by index h). However, in this work, we focus on single phase, polycrystalline, 3D RVEs. Consequently, the physical descriptors that define the material local state in each voxel need to include information about the crystal lattice orientation (i.e., g) which can take values from a continuous domain (i.e., SO(3)) [47]. Therefore, it becomes necessary to modify the representation of the microstructure function from m_s^h to $m_s(g)$ [11]. As a result, the microstructure function $m_s(g)$ prescribes the probability density associated with finding the orientation g in the spatial voxel s [48]. In other words, $m_s(g)$ is essentially the spatially resolved orientation distribution function for voxel s. This function needs to be converted into a set of discrete features. In the present work, this is accomplished using the well-established GSH basis.

Generalized Spherical Harmonics. GSH are known to provide a complete Fourier basis for representations of functions defined over the orientation space. They are expressed mathematically as [42]

$$T_l^{\mu n}(g) = T_l^{\mu n}(\varphi_1, \Phi, \varphi_2) = e^{i\mu\varphi_1} P_l^{\mu n}(\Phi) e^{i\mu\varphi_2} \qquad (1)$$

where $P_l^{\mu n}(\Phi)$ are the associated Legendre polynomials [8,11,31,42]. Furthermore, the GSH basis can be customized to reflect any combinations of crystal and sample symmetries by identifying suitable linear combinations of the unsymmetrized basis [42]. In this work, we will focus on crystalline materials that exhibit cubic-triclinic symmetry (i.e., the crystals exhibit cubic symmetry and the sample exhibits no additional symmetry). Let $\ddot{T}_l^{\mu n}(g)$ denote the set of GSH basis functions that account for the desired symmetry (i.e., cubic-triclinic). This set of symmetrized GSH basis functions is expressed as

$$\ddot{T}_l^{\mu n}(g) = \sum_{\mu=-l}^{\mu=+l} \ddot{A}_l^{\mu n} T_l^{\mu n}(g) \qquad (2)$$

where $\ddot{A}_l^{\mu n}$ are solved to impart the desired crystal and sample symmetries to the new basis. Using the set of symmetrized GSH basis, $\ddot{T}_l^{\mu n}(g)$, we can establish a suitable Fourier representation for the microstructure function as

$$m_s(g) = \sum_{\mu,n,l} K_{ls}^{\mu n} \ddot{T}_l^{\mu n}(g) \qquad (3)$$

where $K_{ls}^{\mu n}$ denotes the GSH coefficients. Since $m_s(g)$ is a real-valued function, it can be shown that half of $K_{ls}^{\mu n}$ are complex conjugates of the other half [42]. Furthermore, for the specific case where only one crystal orientation g_0 is present at voxel s (i.e., $m_s(g) = \delta_s(g - g_0)$), the GSH coefficients are simply obtained as

$$K_{ls}^{\mu n} = \frac{1}{(2l+1)} \ddot{T}_l^{\mu n*}(g_0) \qquad (4)$$

where $\ddot{T}_l^{\mu n*}(g)$ is the complex conjugate of $\ddot{T}_l^{\mu n}(g)$. As a result of these observations, it is possible to map the $K_{ls}^{\mu n}$ into a set of real-valued coefficients, denoted here as $F_{ls}^{\mu n}$. In order to obtain a set of real-valued coefficients from $K_{ls}^{\mu n}$, we take advantage of the fact that the GSH coefficients corresponding to $\mu < 0$ are the complex conjugates of the GSH coefficients corresponding to $\mu > 0$. Furthermore, the GSH coefficients corresponding to $\mu = 0$ are all real. Therefore, saving the real and the imaginary parts of the GSH coefficients corresponding to $\mu < 0$ as two separate real-valued variables in addition to the real-valued GSH coefficients for $\mu = 0$ provides a set of real-valued coefficients that can be uniquely mapped back to the original complex-valued GSH coefficients. The set of real-valued coefficients obtained in this manner

are denoted here as F_s^L. In this simplified notation [31,34], L enumerates all the distinct combinations of μ, n, and l.

The GSH coefficients obtained from Eq. (2) produce an infinite set of local features in each voxel. In the present application, our interest lies in utilizing these features to represent the single crystal orientations present in each voxel. It is well known that only three independent parameters (e.g., Bunge–Euler angles) are needed to fully describe a single crystal orientation. Therefore, the GSH coefficients obtained from Eq. (2) for single crystal orientation distributions in each voxel are indeed related to each other. However, the inter-relationships between the GSH coefficients would be highly nonlinear. As a result, these coefficients can serve as separate input channels to a CNN. In this paper, we will employ a suitably truncated set of real-valued and transformed GSH coefficients to represent the information about the single crystal orientation present within each voxel of the RVE. One of the main advantages of this representation is that the F_s^L coefficients inherently account for all the symmetries and equivalencies associated with cubic crystals. Most importantly, the values of F_s^L provide meaningful measures of the closeness of any selected sets of orientations through the use of simple Euclidean distances between their GSH representations. In fact, GSH-based representations of polycrystalline microstructures have been shown to serve as excellent predictors of their effective properties [32,46,49–51].

The benefits of GSH representations are illustrated through a specific example in Fig. 1(a), which shows two different orientations denoted as A and B, respectively, within the fundamental zone of the orientation space for cubic-triclinic symmetry expressed with Bunge–Euler angles [8,47]. However, in the Bunge–Euler angle representations, they appear very far from each other. In reality, the misorientation between these orientations is only one degree. The GSH-based representation of the corresponding single crystal ODFs, shown in Fig. 1(b), shows them to be indeed very close to each other (with only a small difference in the higher terms of the expansion, F_s^L).

Material Knowledge System Localization Framework. In recent years, our research group has established a computational framework called MKS capable of accurately predicting the local response of a macroscopically imposed tensor (i.e., localization) for polycrystalline microstructures [11,31,34]. This framework, which is based on the statistical continuum theories developed by Kröner [19,20], computes the local response using a series expansion. Each term of the series is determined by a convolution of a physics-capturing microstructure-independent kernel (also called MKS localization kernel) over a suitable higher-order representation of the microstructure field [8,35]. As a result, the MKS framework enables one to obtain the localized response, P_s, of a macroscopically imposed quantity, $\langle P \rangle$, as follows [34]:

$$P_s = \left(\sum_t^S \int_{FZ} \alpha_t(g) m_{s+t}(g) dg + \sum_t^S \sum_{t'}^{S'} \int_{FZ} \int_{FZ} \right.$$
$$\left. \alpha_{tt'}(g, g') m_{s+t}(g) m_{s+t+t'}(g, g') dg' dg + \cdots \right) \langle P \rangle \qquad (5)$$

where $m_s(g)$ is the voxel-based microstructure function introduced earlier, $\alpha_t(g)$ denote the first-order localization kernels, and $\alpha_{tt'}(g, g')$ are the second-order localization kernels. Equation (5) was designed to capture systematically the contributions from all of the local states present in the neighborhood of the spatial voxel of interest identified by s on the response of interest at that same voxel. Previous work has also demonstrated that the systematic introduction of the higher-order terms in the series expansion increases the accuracy of the localization models [36]. However, the incorporation of the higher-order terms is an arduous and complex task, mainly because the number of fit parameters grows significantly [36]. Consequently, extension of the MKS framework to include the higher-order terms becomes
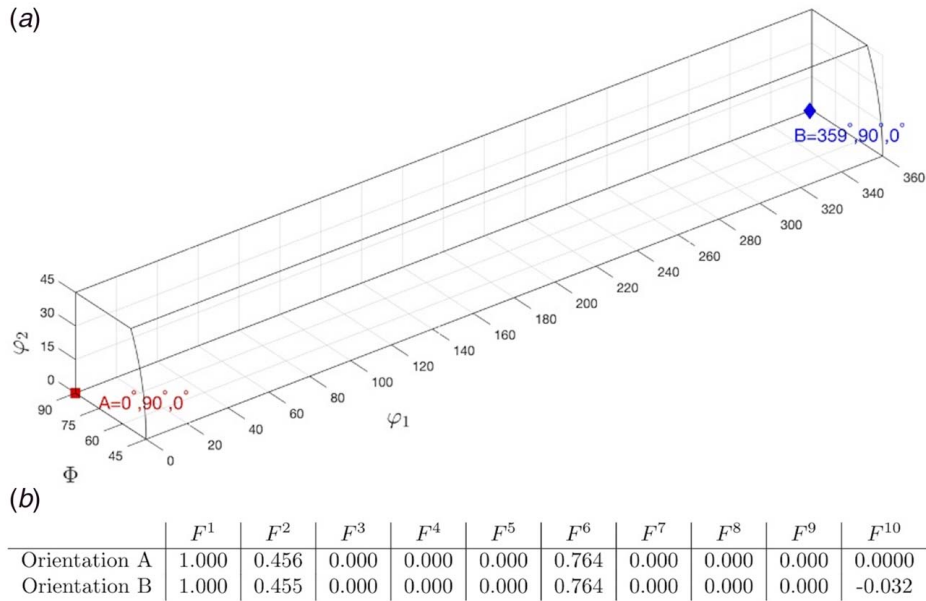
Fig. 1 Comparison of Bunge–Euler angles-based representation versus GSH-based representation of the crystal orientations in a single voxel. (a) Representation of two similar crystal orientations within the fundamental zone for cubic crystals in the Bunge–Euler angle space. (b) Transformed GSH-based representations of the same two crystal orientations from (a) with a truncation level up to $l = 4$ (i.e., $L = 10$).

largely intractable. Recent work has demonstrated the benefits of establishing accurate surrogate localization models using CNNs [37], especially for material systems that exhibit a high contrast in the elastic properties exhibited by the different material local states present in the microstructure. For the plastic deformation of polycrystalline materials, the relevant contrast would be determined by the differences in the plastic properties of the differently oriented grains present in the polycrystal microstructure. Building on our prior work [34], we explore here the viability and benefits of employing CNNs to establishing low-computational cost surrogate models for localization of plastic strain rate tensor in polycrystalline microstructures.

Convolutional Neural Network Model. CNNs are a type of deep learning model used to learn the functional linkage between selected input fields and an output, in a supervised learning setting. These tools have been formulated and utilized to great success in the field of computer vision for diverse tasks such as image classification, object localization, and image segmentation [28,29,52,53]. Recently, CNN models have also been successfully applied beyond the domain of digital image processing in fields such as radio astronomy, medical imaging, bioinformatics, genomics, and materials science [37,38,54–59]. CNNs, akin to other neural network type models, receive an input (e.g., an image or a micrograph) and transform it through a series of different layers to generate a scalar or vector output. CNNs contain at least one or more convolutional layer, where the input is transformed through convolutions with a set of learnt filters, followed by application of a nonlinear function such as sigmoid or rectified linear unit (ReLU) [60] on the output. A key characteristic of this layer is that the filters that are convolved are learnt to identify outputs corresponding to the salient features relevant to the modeling task at hand. This is accomplished by calibrating the weights in each filter to the available training data. Additionally, CNN models may also contain other types of layers designed for dimensionality reduction (e.g., pooling layer) [28,29], inner products (fully connected layer) [29], and normalization [29]. The number and type of the constituent layers, as well as the sequence in which they are applied, are generally referred as the architecture of the CNN model. In the simplest CNN architecture, also known as a feed-

forward network, the layers are sequentially applied to the input data. In other words, the output of one layer is used as the input to the subsequent layer. By modifying the model architecture, one can refine and optimize the learning capacity of a CNN model for a selected modeling task. This lends immense versatility to the CNN modeling framework. As a result, CNN models have widespread adoption and application across a broad spectrum of disciplines [61–64].

In typical supervised learning settings, the model fit parameters are inferred from available training dataset (includes sets of inputs and their corresponding outputs) by minimizing a suitably defined loss (error) function. In CNN implementations, the optimal values of the model fit parameters are determined through an iterative, non-convex optimization algorithm known as stochastic gradient descent [65]. The training of the CNN ends once the loss saturates or it reaches below a pre-established threshold.

Finite Element Models for Generating Training and Validation Data. The CNN models developed in this work will be trained with results from crystal plasticity finite element simulations method (CPFEM) performed on a set of synthetically generated microstructures, which we will refer to as the training set. Subsequently, the calibrated CNN models will be critically validated by comparing their predictions for a different set of microstructures, which we will refer to as the test set, with the corresponding predictions from CPFEM simulations. Note that the test set microstructures are not exposed to the CNNs in their training phase. It is desirable for both the training and test sets to encompass a rich and diverse collections of polycrystalline microstructures. Following prior work [34], we generated the microstructures needed for this work using DREAM.3D software [66]. This tool offers tremendous flexibility in creating a variety of polycrystalline microstructures by changing the grain size, shape, and orientation distributions. The microstructures generated are voxelized microstructures of size $21 \times 21 \times 21$. Previous work has demonstrated that this discretization level is sufficient to train accurate localization models for polycrystalline RVEs [30,34].

DREAM.3D generates voxelized microstructures by packing the RVE with ellipsoidal grains, and then expanding the grains to fill out the RVE. In this study, we generated RVEs with different

grain sizes and two different grain morphologies. The grain size was controlled mainly by changing the targeted number of grains per RVE (i.e., as more grains were packed into an RVE of fixed size, $21 \times 21 \times 21$, the grains became smaller). The grain shapes were assigned as ratios of the axes lengths of the ellipsoids. These ratios were defined as 1:1:1 for the equi-axed grain and 1:1:K (where $K > 1$) for the elongated grains. For the elongated grains, the long direction was varied among the three reference directions of the RVE cube. Table 1 summarizes the different parameters used for generating a total of 21,700 RVEs for the training set and a total of 6000 RVEs for the test set. It is noted here that targeted grain sizes and the assigned orientations to the grains in the test set are intentionally kept distinct from those used in the training set. The orientations of the grains in the training RVEs were selected randomly from a set of 20,000 orientations distributed uniformly in the cubic fundamental zone (see Fig. 1(a)), while the orientations for the testing set were selected similarly from another set of 3500 orientations distributed in the same cubic fundamental zone. In other words, all of the voxels associated with a grain within an RVE are assigned an orientation randomly selected from a set of 20,000 orientations (for the RVEs of the training set) or from a different set of 3500 orientations (for the RVEs of the test set). As a result, the generated RVEs span a wide variety of textures (i.e., orientation distribution functions) as well as intensities because the RVEs with less grains will show stronger textures than the RVEs with a higher number of grains. Therefore, the dataset generated for this work is quite diverse in terms of grain morphology as well as texture. It is important to note that the training set is purposely generated to span larger domains; the test set therefore represents an interpolation of the training set. Figures 2(a) and 2(b) show examples of RVEs from the training set and the test set, respectively.

The kinematic variable of interest to the present work is the velocity gradient tensor, usually denoted by L in the literature. In elastoplastic deformations, L can be decomposed into elastic (L^*) and plastic (L^p) components. Since the focus here is centered on the localization of the plastic deformation experienced by polycrystalline RVEs, L^p is the main quantity of interest. In crystal plasticity models, L^p is expressed as a sum of shearing rates on available slips systems [22]. The symmetric part of L^p is denoted by D^p and is commonly referred as the plastic stretching tensor. D^p captures the information on the plastic strain rate. The anti-symmetric part of L^p is denoted by W^p and is commonly referred as the plastic
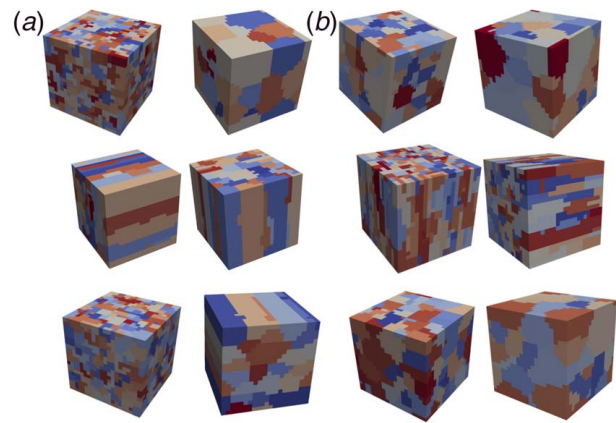

Fig. 2 Selected example RVEs used in this work. (a) Example RVEs from the training set and (b) example RVEs from the test set.

spin tensor. The microscale distribution of L^p (i.e., variation between individual voxels of the polycrystalline RVE) was computed using previously established CPFEM simulations using the commercial FE software ABAQUS [67]. In these models, the elastic-viscoplastic crystal plasticity constitutive laws are implemented at each integration point in the FE mesh with a user material subroutine (UMAT). The specific UMAT used in this study has been extensively validated for fcc metals in prior work [22,68]. The reader is referred to Ref. [22] for a detailed description of the crystal plasticity model used in this work and its implementation in ABAQUS. Furthermore, the CPFEM simulations performed in this work have mostly used the same material parameters described in that earlier work. The main difference is that we simplified the material description so that we can produce accurate training data for the localization of the plastic velocity gradient tensor. Specifically, we employed high values of the elastic stiffness parameters in order to limit the elastic strains to small numbers (<0.001). We also avoided strain hardening hardening by setting all slip hardening rates to zero. The main material parameters used are a slip rate-sensitivity (usually denoted by m) of 0.01 and a reference slip rate (usually denoted by $\dot{\varepsilon}_o$) of 0.001 s^{-1} [22]. For the computations described here, the local values of L^p in each voxel of the RVE are not affected by the value of the initial slip resistance (usually denoted by s_o), which is assumed to be the same for all slip systems in all grains in the RVE.

Subsequently, the voxelized polycrystalline RVEs in both the training and test sets were converted to FE meshes, where each voxel became an element of the FE mesh. Cubic continuum elements with eight nodes (C3D8) [67] were used. Each RVE was subjected to periodic boundary conditions corresponding to the following macroscopic plastic velocity gradient [33]:

$$\langle L^p \rangle = \begin{bmatrix} 0.4 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & -0.8 \end{bmatrix} \times 0.02 \text{s}^{-1} \qquad (6)$$

The CPFEM simulations provided the ground-truth data on how the imposed plastic velocity gradient tensor in Eq. (6) is partitioned into the individual elements (i.e., voxels) of the RVE. For each CPFEM simulation, we imposed a strain of 0.02 over a time period of 1 s. Finally, in order to completely avoid any possible effects of the drastically reduced elastic strains, the averaged plastic velocity gradient tensor was computed in each voxel over the 0.01–0.02 range of the total imposed strain.

The averaging was performed over the integration points of the element (i.e., voxel) as well as the different time-steps involved in the specified strain range. Each voxel of the FE simulation provided one data point for the localization of polycrystal plasticity. In other words, each FE simulation performed generated a set of 9261 data

Table 1 Parameters used for generating the digital polycrystalline RVEs used in this work

Target number of grains	K	Orientation of the longest axis of the grains	Number of RVEs	Purpose
22	1	X	4000	Training
100	1	X	4000	Training
162	1	X	500	Training
347	1	X	1200	Training
860	1	X	1000	Training
3500	1	X	500	Training
7280	1	X	1500	Training
22	10	X	1000	Training
22	10	Y	1000	Training
22	10	Z	1000	Training
35	10	X	1333	Training
35	10	Y	1333	Training
35	10	Z	1334	Training
101	10	X	666	Training
101	10	Y	666	Training
101	10	Z	668	Training
45	1	X	3000	Test
127	1	X	1000	Test
570	5	X	1000	Test
1300	3	Y	1000	Test

points, which were used either in the training set or the test set (corresponding to which RVEs were used in the FE simulation). It should be noted that in order to generate the training data for the CNN-based localization surrogate model, the CPFEM simulations performed in this work only needed small strain levels. Nevertheless, it is important to recognize that in order to extend the surrogate model predictions to larger strains, an efficient time-integration procedure would also be needed. Such a time-integration procedure still needs to be developed and is beyond the scope of the present work.

The 27,700 FE simulations performed for this study incurred significant computational cost. Each CPFEM simulation took approximately 90 min on a single 2.4 GHz AMD processor node in the Georgia Tech super computer cluster. The central benefit of the CNN-based surrogate models developed in this work is that the training represents a one-time cost. All test evaluations from the CNN models for new microstructures are very cheap (∼0.1 s for each new polycrystalline RVE).

Convolutional Neural Networks for the Localization of Polycrystal Plasticity

The primary goal of this work is to critically evaluate the benefits of using GSH representations (see Fig. 1(b)) versus the Bunge–Euler angles representation (see Fig. 1(a)) as the input for a CNN-based surrogate model. Towards this goal, in one implementation, we will directly use the Bunge–Euler angles expressed in the fundamental zone of the orientation space to train a CNN surrogate model for predicting the local plastic velocity gradient tensor in each voxel of the polycrystalline RVE. For the second implementation, we will use the GSH-based representations of the spatially resolved ODF for each voxel, truncated to different numbers of terms in the Fourier series, as different input channels for the CNN surrogate model.

The CNN surrogate models developed in this work are trained to predict the local (average over a single voxel) values of the plastic velocity gradient tensor in 3D polycrystalline RVEs. The specific input into the CNN will be the 3D neighborhood of the voxel of interest for which the components of the plastic velocity gradient tensor are being predicted. For this work, the input to the CNN was defined as a cubical neighborhood comprising $21 \times 21 \times 21$ voxels centered around the voxel of interest. The given polycrystalline RVE (also of size $21 \times 21 \times 21$) was extended assuming periodic boundaries, consistent with the boundary conditions imposed

in the CPFEM simulations of the RVEs described earlier in the "Finite Element Models for Generating Training and Validation Data Section," to define the required neighborhoods.

The CNN models developed in this work implicitly featurize the 3D neighborhoods of individual voxels in the polycrystalline structures using 3D convolutional layers. A 3D convolutional layer comprises multiple sets of filters. Each set of filters in each convolution layer is of size $q \times p \times p \times p$, where p denotes the number of voxels in each spatial direction of a 3D filter and q denotes the number of channels. The number of channels is dictated by the number of 3D inputs to the specific convolution layer. As a specific example, for the first convolutional layer, q would correspond to the number of input channels to the CNN. This means $q = 3$ for the Bunge–Euler angle implementation and $q = L - 1$ for the GSH implementation (this is because the first GSH coefficient is always equal to one irrespective of the grain orientation in the voxel). Sequentially stacking 3D convolutional layers enables more intricate featurizations by identifying higher-order spatial correlations within the input images [38]. In addition, incorporating ReLU (defined as $f = \max(0, x)$) function between the convolutional layer allows for the identification of nonlinear features. The features identified by the convolutional layers are then mapped to the targets, i.e., the nine components of the local plastic velocity gradient tensor. This mapping is accomplished using fully connected layers. These fully connected layers perform linear and nonlinear (when coupled with activation functions such as ReLU) mappings between the features identified at the end of the convolutional layers and the selected targets.

In this work, we systematically explored multiple CNN architectures with varying learning capacities in order to identify the architecture that produces the most accurate and robust surrogate localization model for the task at hand. As an example, Fig. 3 schematically depicts how the CNN transforms the polycrystalline neighborhood of one voxel within the RVE into its corresponding nine components of the plastic velocity gradient tensor for one of the feed-forward CNN architectures used in this work for the Bunge–Euler angle implementation.

The convolution operations utilized in the CNN architectures explored in this study were performed without the periodic assumption. Therefore, the spatial domain size of the output of each convolutional layer is smaller compared to the size of the spatial domain of its input. Since the filters used were of size $5 \times 5 \times 5$, the spatial domain size shrunk by two voxels near all boundaries. In order words, the input domain size of $21 \times 21 \times 21$ shrunk to an output
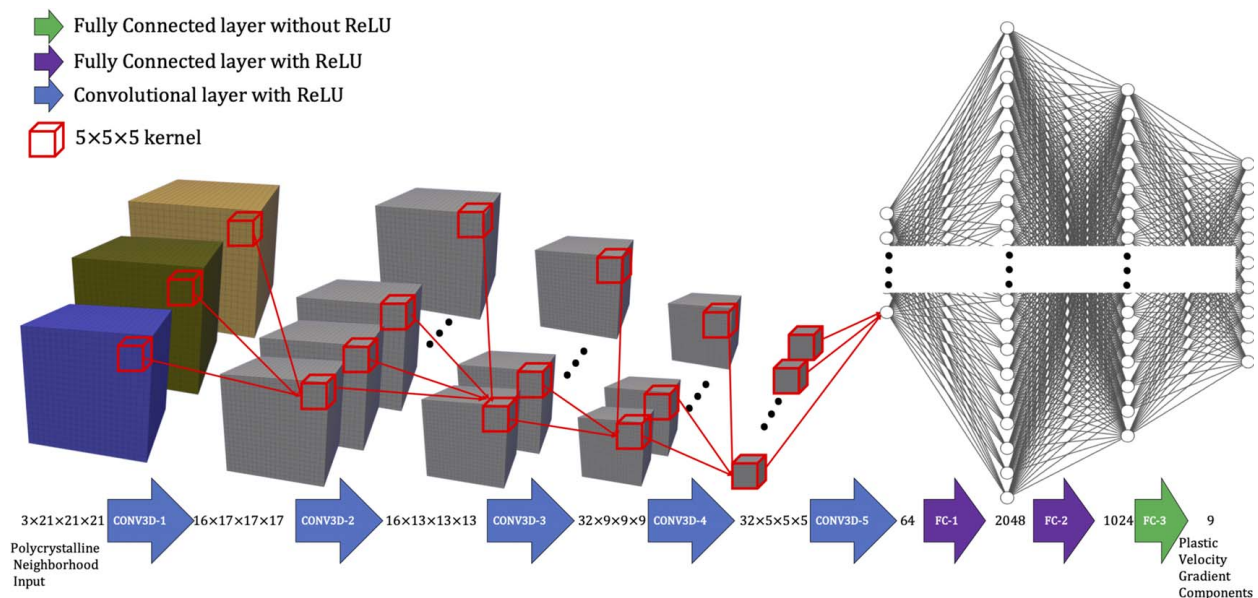

Fig. 3 Data flow (i.e., from inputs to output) through the feed-forward CNN architecture of model A (see Table 2) for the Bunge–Euler representation of the crystal orientation in each voxel

Table 2 The four different CNN architectures evaluated in this work

Layers	Model A	Model B	Model C	Model D
Input	$q \times 21 \times 21 \times 21$	$q \times 21 \times 21 \times 21$	$q \times 21 \times 21 \times 21$	$q \times 21 \times 21 \times 21$
Conv3D-1	$16 \times q \times 5 \times 5 \times 5$	$32 \times q \times 5 \times 5 \times 5$	$64 \times q \times 5 \times 5 \times 5$	$128 \times q \times 5 \times 5 \times 5$
Conv3D-2	$16 \times 16 \times 5 \times 5 \times 5$	$32 \times 32 \times 5 \times 5 \times 5$	$64 \times 64 \times 5 \times 5 \times 5$	$128 \times 128 \times 5 \times 5 \times 5$
Conv3D-3	$32 \times 16 \times 5 \times 5 \times 5$	$64 \times 32 \times 5 \times 5 \times 5$	$128 \times 64 \times 5 \times 5$	$256 \times 128 \times 5 \times 5 \times 5$
Conv3D-4	$32 \times 32 \times 5 \times 5 \times 5$	$64 \times 64 \times 5 \times 5 \times 5$	$128 \times 128 \times 5 \times 5 \times 5$	$256 \times 256 \times 5 \times 5 \times 5$
Conv3D-5	$64 \times 32 \times 5 \times 5 \times 5$	$128 \times 64 \times 5 \times 5 \times 5$	$256 \times 128 \times 5 \times 5 \times 5$	$512 \times 256 \times 5 \times 5 \times 5$
FC-1	2048×64	2048×512	2048×256	2048×512
FC-2	1024×2048	1024×2048	1024×2048	1024×2048
FC-3	9×1024	9×1024	9×1024	9×1024

domain size of $17 \times 17 \times 17$ at the end of the first convolutional layer. This shrinking continued with each convolutional layer and produced single features (i.e., an output domain of $1 \times 1 \times 1$) after five convolutional layers. Additionally, it should be noted that the number of distinct filters applied in each convolution layer is determined by the product of the number of input channels and the number of output channels for the layer. As an example, in the first convolutional layer shown in the CNN architecture depicted in Fig. 3, the number of distinct filters (each of size $5 \times 5 \times 5$) is $3 \times 16 = 48$, while it is $16 \times 16 = 256$ for the second convolutional layer. In the CNN architecture shown in Fig. 3, we obtain 64 features after applying five 3D convolutional layers (along with the associated ReLUs). Subsequently, three fully connected layers are applied on the 64 features to produce the desired nonlinear mappings to the nine targets (i.e., the components of the plastic velocity gradient tensor). Note that a ReLU is not applied in the last layer because the values obtained in this layer represent the targets. The use of three fully connected regression layers for such mappings is common in the deep learning domain [69]. The decision to use the type of architecture shown in Fig. 3 is largely motivated by prior work on MKS localization [31,34].

In this work, we have considered four different variations of the CNN architecture shown in Fig. 3. The convolution filter size (i.e., $p = 5$) and the last three fully connected layers are kept the same in all the variations. The learning capacity of the different CNN models was modulated by varying the numbers of filters in the five 3D convolutional layers. Table 2 summarizes the details of the four different CNN models, labeled A through D, explored in this work. Note that the size of the first layer is different for the different representations of the grain orientations explored in this work. As already mentioned, $q = 3$ for the Bunge–Euler representation. For the truncated GSH representations, the values of q corresponded to truncation levels of $l = 4$, 6, 8 (see Eq. (1)), which resulted in values of 9, 22, and 39, respectively.

Loss Function for the Convolutional Neural Network Model. The parameters in the CNN model (i.e., the values of the different weights and biases defining the convolution filters) are inferred during the training process using back-propagation [70] coupled with a stochastic optimization algorithm known as ADAM [71].

This optimization technique aims to minimize a loss function that quantifies the deviation of the model predictions from the known ground-truth values in the training set (obtained from the CPFEM simulations). The loss employed in this work is defined as

$$\text{loss}\% = \frac{1}{9 \times S \times |\langle L^p \rangle|} \sum_{u=1}^{3} \sum_{v=1}^{3} \sum_{s=1}^{S}$$
$$\left| L_{uv}^p[s]_{\text{CPFEM}} - L_{uv}^p[s]_{\text{CNN}} \right| \times 100\% \qquad (7)$$

Results and Discussion

The four different architectures presented in Table 2 along with the two different representations of the crystal lattice orientation

were critically evaluated for their fidelities in predicting the local values (in each voxel) of the plastic velocity gradient tensor. This task therefore included a total of four CNN models using the Bunge–Euler representation and 12 CNN models using the GSH representations corresponding to the three truncation levels identified earlier. Each CNN model was trained for 15 epochs (each epoch here represents one update of the CNN model using the full training set) at a constant learning rate of 1×10^{-5} using the ADAM optimizer [71] using four V100 NVIDIA GPU cores on Georgia Tech's super computer cluster. Each epoch took between 20 and 100 min for the different models.

The performance of each trained CNN model was assessed by computing the loss defined in Eq. (7) for the test set. Note that the evaluation of each test RVE by each CNN model produces one value of loss. Figure 4 presents the distribution of the loss values for the test RVEs for the different CNN architectures and the different representations of the crystal orientations using violin plots. These plots provide information about the median (denoted by the white dot within the black box) as well as the interquartile range (which are the upper and lower limits of the black box). Furthermore, violin plots present information about the probability distribution of the data in a clear and concise manner since wider sections of the violin represent a higher probability that members of the population will take on the given value. In a similar manner, the thinner sections represent lower probability values. Figure 4 shows that model D produced the best performance. It is also seen that the CNN models using GSH representations of the crystal orientation consistently performed better than the CNN model using the Bunge–Euler representation. This observation clearly supports the central hypothesis of this work that GSH representations offer an improved representation of the crystal orientations when feeding polycrystalline RVEs as inputs to CNN models.

Model D with the GSH truncation level of $l = 4$ was identified as the best performing model among all the CNN models explored in this work. The improvement in accuracy obtained by increasing the GSH truncation level from $l = 4$ to $l = 6$ and further to $l = 8$ is minimal. This observation implies that the features identified after the five convolutional layers are not further refined significantly with the increasing levels of GSH truncation. This is not surprising given that the GSH coefficients of single crystal orientations are not independent. In other words, the CNN models explored here are able to learn all of the salient nonlinear features in the polycrystalline neighborhoods needed to predict the local values of the plastic velocity gradient tensor using only the GSH representations to $l = 4$. This CNN model was further trained in additional epochs (beyond the initial 15 epochs) using the same optimizer and learning rate until the average loss on the test set exhibited asymptotic convergence. Although the training of the CNN is performed using solely the training set, after every epoch of training the average loss on the test was also calculated in order to monitor the performance improvements with each training epoch. Utilizing this information, the training epochs were stopped when the average loss on the test set did not decrease by greater than 0.1 over two epochs. Using this stopping criteria, the training of model D with GSH
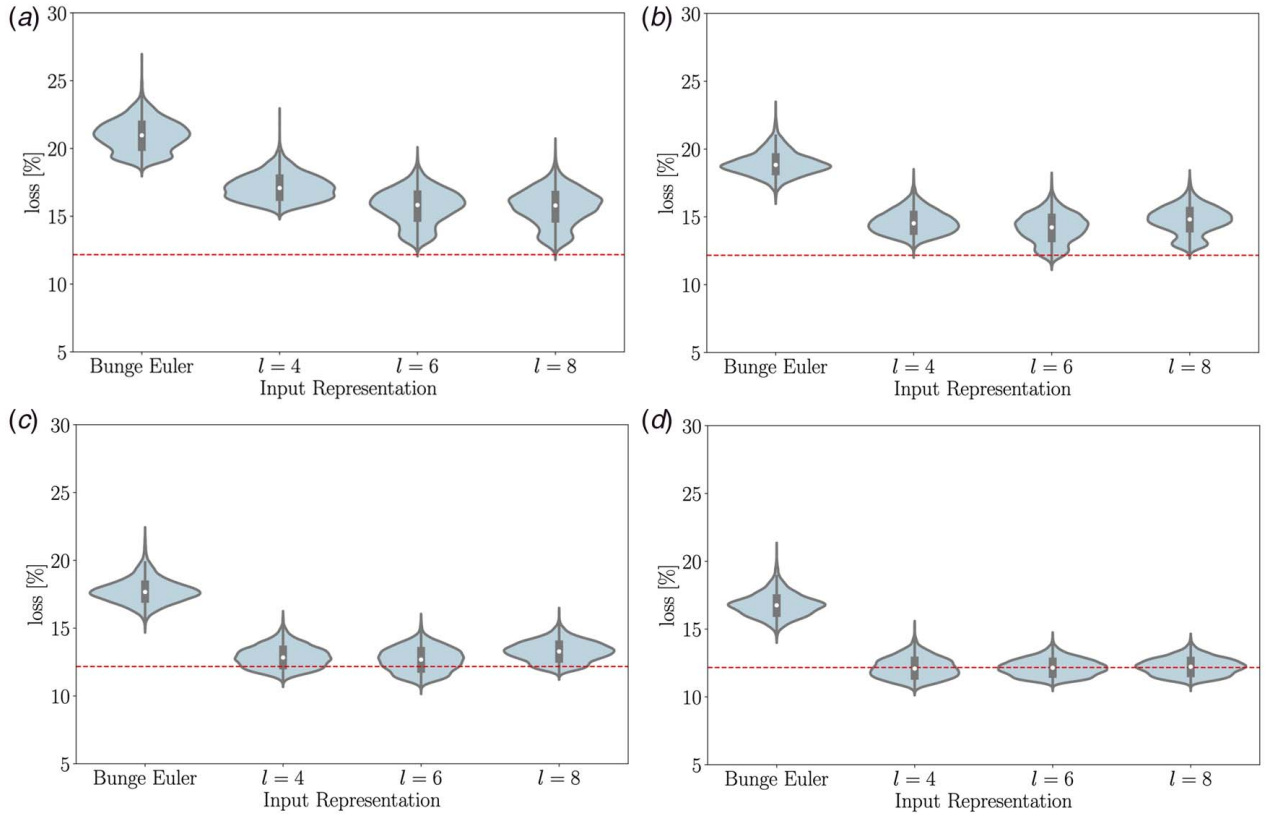
Fig. 4 Distributions of the loss values obtained for the different representations of the internal structure of the polycrystalline RVE using violin plots. (*a*) Model A. (*b*) Model B. (*c*) Model C. (*d*) Model D. The dashed line indicates the average loss on the test set for the best performing model (i.e., model D) and corresponds to a value of 12.25%.

representations to $l = 4$ was stopped at 30 epochs. This improved model produced an average loss of 11.25% on the test set. This loss was quite similar to the loss realized on the training set, indicating that the calibrated CNN model has not been over-fit to the training data.

Next, the best CNN model identified above was further evaluated for its accuracy in predicting each component of the plastic velocity gradient tensor using the following modified version of the loss metric:

$$\text{Err}_{uv}\% = \frac{1}{S \times |\langle L^p \rangle|} \sum_{s=1}^{S} \left| L_{uv}^p[s]_{\text{CPFEM}} - L_{uv}^p[s]_{\text{CNN}} \right| \times 100\% \quad (8)$$

The loss metric defined above was used to compare the accuracy of the CNN model produced here against the previous benchmark, which was taken to be the first-order MKS localization model established in prior work [34]. The benchmark MKS model was also calibrated on the exact same training set. It is noted that the benchmark MKS model (i.e., the best MKS-based localization model identified in prior work [34]) used a GSH truncation to $l = 8$. Figure 5 presents the test set loss distributions for each individual component of the plastic velocity gradient tensor (obtained using Eq. (8)) for the best CNN model produced in this work as well as the benchmark MKS model. Clearly, the CNN model outperforms the benchmark MKS model for all components of the plastic velocity gradient tensor. More specifically, it is seen that the improvements are the greatest for the off-diagonal components of the plastic velocity gradient tensor.

In order to further analyze and understand the performance of the CNN model in the predictions of the plastic velocity gradient tensor components, we also investigated the individual losses in its symmetric and anti-symmetric components. Figures 6(*a*)–6(*c*) show

an example comparison of a selected symmetric component of the plastic velocity gradient tensor (i.e., the (1, 3) component of the plastic stretching tensor, D^p) predicted by the CNN model, the benchmark MKS model, and the ground-truth CPFEM simulation for a middle section of a randomly selected test RVE. Figure 6(*d*) shows a comparison of the probability density distributions of the predicted values for D^p (1, 3) across the entire selected RVE (i.e., for all voxels in the RVE). On the other hand, Figs. 7(*a*)–7(*d*) present the same comparisons for an anti-symmetric component of the plastic velocity gradient tensor (i.e., the (1, 3) component of the plastic spin tensor, W^p). It is clear from these plots that the better performance of the CNN model produced in this work is attributable to its improved predictions of the anti-symmetric components of the plastic velocity gradient tensor. More specifically, the CNN model is able to capture better the extreme values of the plastic spin tensor components (the same observation is also made, although to a much lesser extent, for the extreme values of the plastic stretching tensor components). It should be noted that the plastic stretching tensor fields are controlled directly by the equilibrium-satisfying stress fields, whereas the plastic spin tensor fields are only indirectly controlled by the compatibility requirements on the overall displacement fields. Therefore, the plastic spin tensor fields can be expected to exhibit much more complex variations within a heterogeneous RVE. Therefore, it is quite reasonable that one needs a much better set of nonlinear features in order to predict accurately the components of the plastic spin tensor compared to the components of the plastic stretching tensor. The comparisons shown in Figs. 6 and 7 suggest that the CNN model produced in this work is able to accomplish this task significantly better than the benchmark MKS model.

The performances of model D established using Bunge–Euler representations are compared with model D established using the GSH truncation level of $l = 4$ in Fig. 8 for the individual components of the plastic velocity gradient tensor in the test set (using
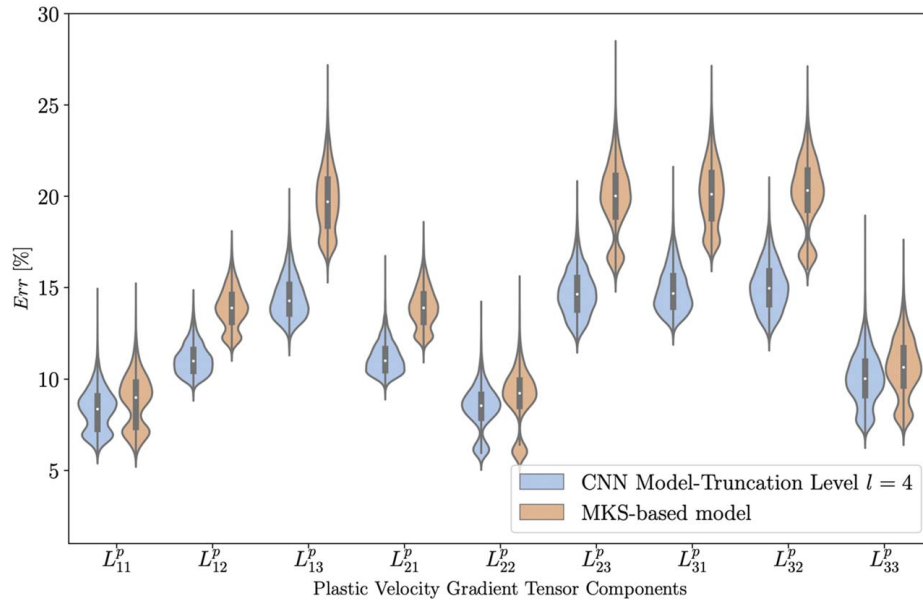
Fig. 5 Loss distributions for each component of the plastic velocity gradient tensor obtained for the best CNN model produced in this work and for the benchmark MKS model

Eq. (8)). The model built using the Bunge–Euler representation was trained for 60 epochs, while the one using the GSH representations needed only 30 epochs. Clearly, the CNN model using the GSH representation for the polycrystalline neighborhood consistently outperforms the CNN model using the Bunge–Euler representation for all components of the target. The better performance at a smaller

number of epochs clearly establishes that the GSH representations produce better features for the localization of the plastic velocity gradient tensor in polycrystalline RVEs.

It is important to recognize the computational savings obtained using a CNN-based localization surrogate model. Once trained, a CNN-based model incurs a minimal computational cost to predict
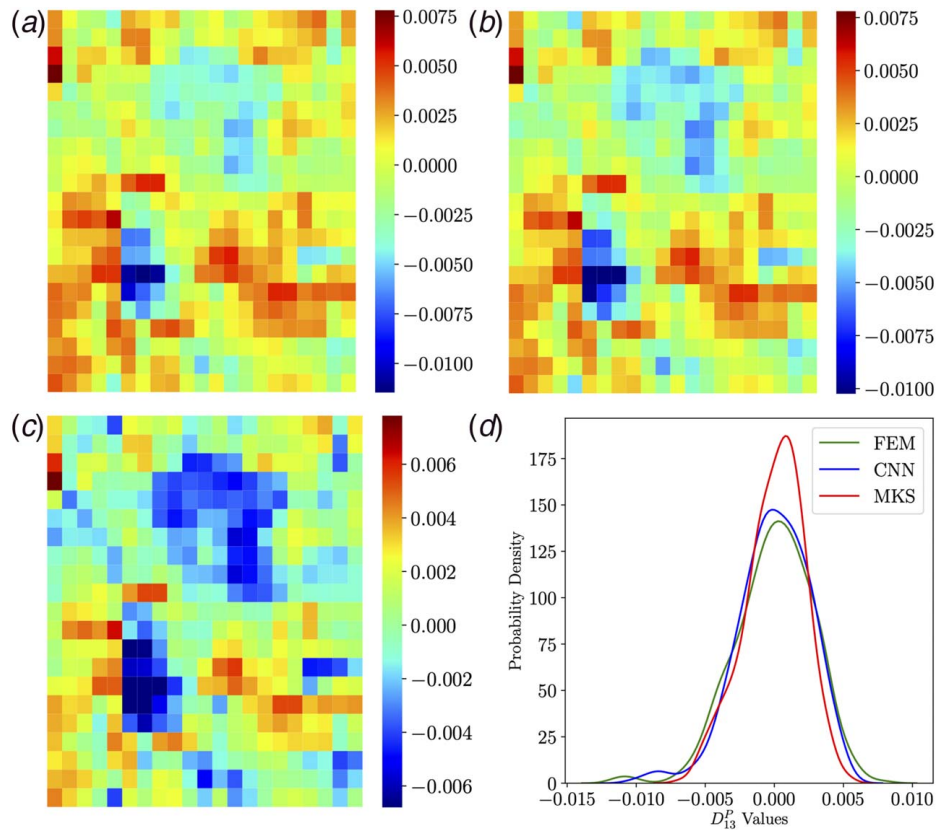

Fig. 6 Visualization and comparison of the true and predicted D^p_{13} values. (a) CPFEM obtained values of D^p_{13} ($D^p_{13CPFEM}$). (b) CNN obtained values of D^p_{13} (D^p_{13CNN}). (c) MKS obtained values of D^p_{13} (D^p_{13MKS}). (d) Probability density distribution of D^p_{13} for the different models and the ground truth.
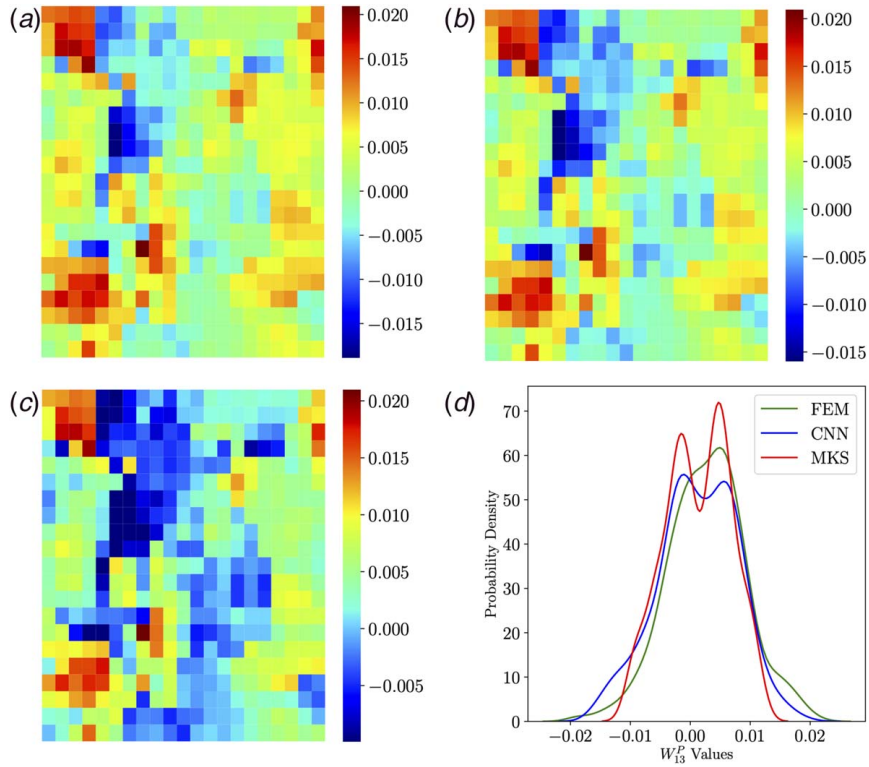
Fig. 7 Visualization and comparison of the true and predicted W_{13}^p values. (a) CPFEM obtained values of W_{13}^p ($W_{13CPFEM}^p$). (b) CNN obtained values of W_{13}^p (W_{13CNN}^p). (c) MKS obtained values of W_{13}^p (W_{13MKS}^p). (d) Probability density distribution of W_{13}^p for the different models and the ground truth.
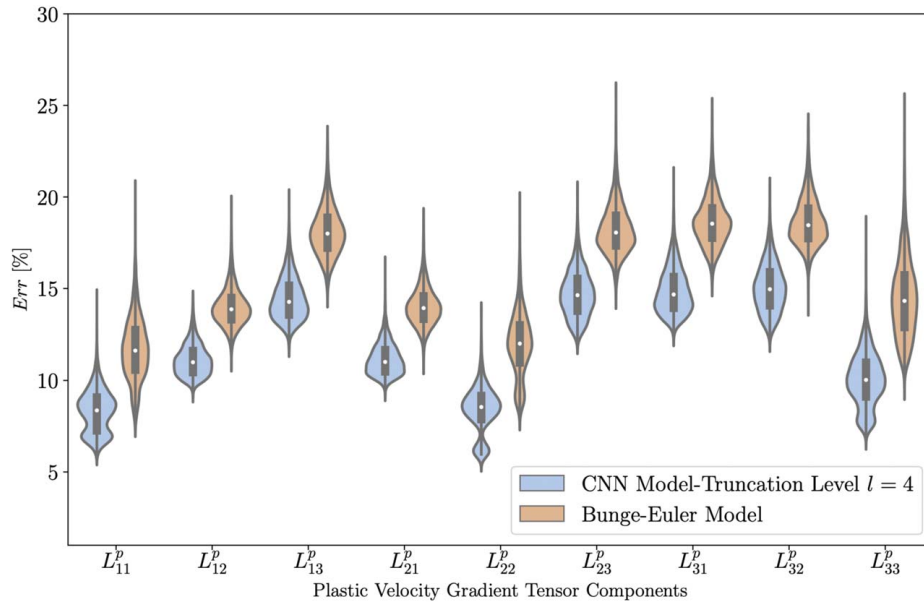

Fig. 8 Comparison of the loss distributions model D obtained with the Bunge–Euler representations and the GSH representations to $l = 4$

the local responses for new/unseen RVEs. As an example, model D established using the GSH truncation level of $l = 4$ predicts the local values of the plastic velocity gradient tensors for a macroscopically imposed tensor in 0.159 s using four NVIDIA V100 GPU cores. Therefore, the computational savings obtained are quite significant. In order to demonstrate that the computational savings obtained are not caused by the advanced computational resources used by the

model (i.e., the GPU cores), we used the CNN-based model on the exact same computing resource used to perform the traditional CPFEM simulation (which corresponds to a single 2.4 GHz AMD processor node) to predict the local values of the plastic velocity gradient. The CNN computation took 5 s, while the CPFEM computation took 5400 s. Therefore, on the same computational resource, the CNN-based localization surrogate model is able to predict the

local values of the plastic velocity gradient tensor in a polycrystalline RVE about 1000 times faster than the corresponding CPFEM simulation.

Conclusions

Convolutional neural networks offer a powerful emergent toolset for establishing low-computational cost surrogate models for capturing localization in heterogenous material systems. These tools will be invaluable for practical materials design efforts. However, the single crystal orientation in the individual voxels of a polycrystalline RVE needs a suitable representation in order to successfully leverage the CNNs for the localization surrogate models. This paper suggested and demonstrated that the use of GSH representations of the crystal orientation serves as a better input to the CNN compared to the conventionally used Bunge–Euler angle representations. Specifically, it is shown that the GSH representation of the polycrystalline neighborhood outperformed the Bunge–Euler representations for a range of CNN architectures. It was also shown that the best CNN model produced in this work significantly outperformed a benchmark model from prior work. Although the benefits of the GSH representation have been shown for the localization of the plastic velocity gradient tensor, it is expected that the results can be generalized for other types of material responses (e.g., creep, fatigue). This is mainly because the GSH representations of the crystal orientation implicitly reflect and account for the crystal symmetries.

Acknowledgment

Conflict of Interest

There are no conflicts of interest.

References

[1] Roters, F., Eisenlohr, P., Hantcherli, L., Tjahjanto, D. D., Bieler, T. R., and Raabe, D., 2010, "Overview of Constitutive Laws, Kinematics, Homogenization and Multiscale Methods in Crystal Plasticity Finite-Element Modeling: Theory, Experiments, Applications," Acta Mater., **58**(4), pp. 1152–1211.

[2] Bishop, J., and Hill, R., 1951, "Xlvi. A Theory of the Plastic Distortion of a Polycrystalline Aggregate Under Combined Stresses," Lond. Edinb. Dublin Philos. Mag. J. Sci., **42**(327), pp. 414–427.

[3] Bishop, J., and Hill, R., 1951, "Cxxviii. A Theoretical Derivation of the Plastic Properties of a Polycrystalline Face-Centred Metal," Lond. Edinb. Dublin Philos. Mag. J. Sci., **42**(334), pp. 1298–1307.

[4] Kröner, E., 1961, "On the Plastic Deformation of Polycrystals," Acta Metall., **9**(2), pp. 155–161.

[5] McDowell, D. L., Panchal, J., Choi, H. -J., Seepersad, C., Allen, J., and Mistree, F., 2009, *Integrated Design of Multiscale, Multifunctional Materials and Products*, Butterworth-Heinemann, Burlington, MA.

[6] Choi, H.-J., Mcdowell, D. L., Allen, J. K., and Mistree, F., 2008, "An Inductive Design Exploration Method for Hierarchical Systems Design Under Uncertainty," Eng. Optim., **40**(4), pp. 287–307.

[7] McDowell, D. L., and Olson, G. B., 2008, "Concurrent Design of Hierarchical Materials and Structures," *Scientific Modeling and Simulations*, S. Yip, and T. Diaz de la Rubia, eds., Springer, Dordrecht, pp. 207–240.

[8] Adams, B. L., Kalidindi, S., and Fullwood, D. T., 2012, *Microstructure Sensitive Design for Performance Optimization*, Butterworth-Heinemann, Waltham, MA.

[9] Mackenzie, A., Hancock, J., and Brown, D., 1977, "On the Influence of State of Stress on Ductile Failure Initiation in High Strength Steels," Eng. Fract. Mech., **9**(1), pp. 167–188.

[10] Rice, J. R., and Tracey, D. M., 1969, "On the Ductile Enlargement of Voids in Triaxial Stress Fields," J. Mech. Phys. Solids, **17**(3), pp. 201–217.

[11] Kalidindi, S. R., 2015, *Hierarchical Materials Informatics: Novel Analytics for Materials Data*, Elsevier, Waltham, MA.

[12] Adams, B. L., and Olson, T., 1998, "The Mesostructure-Properties Linkage in Polycrystals," Prog. Mater. Sci., **43**(1), pp. 1–87.

[13] Luscher, D. J., McDowell, D. L., and Bronkhorst, C. A., 2010, "A Second Gradient Theoretical Framework for Hierarchical Multiscale Modeling of Materials," Int. J. Plast., **26**(8), pp. 1248–1275.

[14] Milton, G. W., and Sawicki, A., 2003, "Theory of Composites. Cambridge Monographs on Applied and Computational Mathematics," ASME Appl. Mech. Rev., **56**(2), pp. B27–B28.

[15] Castaneda, P. P., 2002, "Second-Order Homogenization Estimates for Nonlinear Composites Incorporating Field Fluctuations: I—Theory," J. Mech. Phys. Solids, **50**(4), pp. 737–757.

[16] McDowell, D., and Dunne, F., 2010, "Microstructure-Sensitive Computational Modeling of Fatigue Crack Formation," Int. J. Fatigue, **32**(9), pp. 1521–1542, Emerging Frontiers in Fatigue.

[17] Fish, J., Yu, Q., and Shek, K., 1999, "Computational Damage Mechanics for Composite Materials Based on Mathematical Homogenization," Int. J. Numer. Methods Eng., **45**(11), pp. 1657–1679.

[18] Qu, J., and Cherkaoui, M., 2006, *Fundamentals of Micromechanics of Solids*, Wiley, Hoboken, NJ.

[19] Kröner, E., 1977, "Bounds for Effective Elastic Moduli of Disordered Materials," J. Mech. Phys. Solids, **25**(2), pp. 137–155.

[20] Kröner, E., 1986, "Statistical Modelling," *Modelling Small Deformations of Polycrystals*, J. Gittus, and J. Zarka, eds., Springer, Heidelberg, pp. 229–291.

[21] Anand, L., and Kothari, M., 1996, "A Computational Procedure for Rate-Independent Crystal Plasticity," J. Mech. Phys. Solids, **44**(4), pp. 525–558.

[22] Kalidindi, S. R., Bronkhorst, C. A., and Anand, L., 1992, "Crystallographic Texture Evolution in Bulk Deformation Processing of fcc Metals," J. Mech. Phys. Solids, **40**(3), pp. 537–569.

[23] Peirce, D., Asaro, R., and Needleman, A., 1982, "An Analysis of Nonuniform and Localized Deformation in Ductile Single Crystals," Acta Metall., **30**(6), pp. 1087–1119.

[24] Asaro, R. J., and Needleman, A., 1985, "Overview No. 42 Texture Development and Strain Hardening in Rate Dependent Polycrystals," Acta Metall., **33**(6), pp. 923–953.

[25] Zbib, H. M., and Diaz de la Rubia, T., 2002, "A Multiscale Model of Plasticity," Int. J. Plast., **18**(9), pp. 1133–1163.

[26] Tan, J., Villa, U., Shamsaei, N., Shao, S., Zbib, H. M., and Faghihi, D., 2021, "A Predictive Discrete-Continuum Multiscale Model of Plasticity With Quantified Uncertainty," Int. J. Plast., **138**(Special Issue on Multi-scale Modeling and Characterization of Material Behavior in Normal and Harsh Environments in Honor of Professor George Z. Voyiadjis), p. 102935.

[27] Chehade, A. A., Belgasam, T. M., Ayoub, G., and Zbib, H. M., 2020, "Accelerating the Discovery of New dp Steel Using Machine Learning-Based Multiscale Materials Simulations," Metall. Mater. Trans. A, **51**(6), pp. 3268–3279.

[28] Lecun, Y., Bengio, Y., and Hinton, G., 2015, "Deep Learning," Nature, **521**(7553), pp. 436–444.

[29] Goodfellow, I, Bengio, Y., Courville, A., and Bengio, Y., 2016, *Deep Learning*, Vol. 1, MIT Press, Cambridge, MA.

[30] Yabansu, Y. C., and Kalidindi, S. R., 2015, "Representation and Calibration of Elastic Localization Kernels for a Broad Class of Cubic Polycrystals," Acta Mater., **94**, pp. 26–35.

[31] Yabansu, Y. C., Patel, D. K., and Kalidindi, S. R., 2014, "Calibrated Localization Relationships for Elastic Response of Polycrystalline Aggregates," Acta Mater., **81**, pp. 151–160.

[32] Paulson, N. H., Priddy, M. W., McDowell, D. L., and Kalidindi, S. R., 2017, "Reduced-Order Structure-Property Linkages for Polycrystalline Microstructures Based on 2-Point Statistics," Acta Mater., **129**, pp. 428–438.

[33] de Oca Zapiain, D. M., Popova, E., and Kalidindi, S. R., 2017, "Prediction of Microscale Plastic Strain Rate Fields in Two-Phase Composites Subjected to an Arbitrary Macroscale Strain Rate Using the Materials Knowledge System Framework," Acta Mater., **141**, pp. 230–240.

[34] de Oca Zapiain, D. M., and Kalidindi, S. R., 2019, "Localization Models for the Plastic Response of Polycrystalline Materials Using the Material Knowledge Systems Framework," Modell. Simul. Mater. Sci. Eng., **27**(7), p. 074008.

[35] Landi, G., Niezgoda, S. R., and Kalidindi, S. R., 2010, "Multi-Scale Modeling of Elastic Response of Three-Dimensional Voxel-Based Microstructure Datasets Using Novel DFT-Based Knowledge Systems," Acta Mater., **58**(7), pp. 2716–2725.

[36] Fast, T., and Kalidindi, S. R., 2011, "Formulation and Calibration of Higher-Order Elastic Localization Relationships Using the MKS Approach," Acta Mater., **59**(11), pp. 4595–4605.

[37] Yang, Z., Yabansu, Y. C., Al-Bahrani, R., Liao, W.-K., Choudhary, A. N., Kalidindi, S. R., and Agrawal, A., 2018, "Deep Learning Approaches for Mining Structure-Property Linkages in High Contrast Composites From Simulation Datasets," Comput. Mater. Sci., **151**, pp. 278–287.

[38] Yang, Z., Yabansu, Y. C., Jha, D., Choudhary, A. N., Kalidindi, S. R., and Agrawal, A., 2019, "Establishing Structure-Property Localization Linkages for Elastic Deformation of Three-Dimensional High Contrast Composites Using Deep Learning Approaches," Acta Mater., **166**, pp. 335–345.

[39] Yang, Z., Al-Bahrani, R., Reid, A. C., Papanikolaou, S., Kalidindi, S. R., Liao, W. K., Choudhary, A., and Agrawal, A., 2019, "Deep Learning Based Domain Knowledge Integration for Small Datasets: Illustrative Applications in Materials

Informatics," Proceedings of the International Joint Conference on Neural Networks, Budapest, Hungary, July 14–19, pp. 1–8.

[40] Taylor, G. I., 1938, "Plastic Strain in Metals," J. Inst. Metals, **62**, pp. 307–324.

[41] Sachs, G., 1929, "Zur ableitung einer fliessbedingung," *Messages From the German Material Testing Institutes*, O. Bauer, M. Hansen, F. V. Göler, G. Sachs, E. Schmid, G. Wassermann, K. Sipp, H. Sieglerschmidt, R. Karnop, W. Kuntze, K. Lute, R. Eisenschitz, B. Rabinowitsch, K. Weissenberg, W. Boas, and M. Masima, eds., Springer, Berlin, pp. 94–97.

[42] Bunge, H.-J., 2013, *Texture Analysis in Materials Science: Mathematical Methods*, Elsevier, Waltham, MA.

[43] Miyazawa, Y., Briffod, F., Shiraiwa, T., and Enoki, M., 2019, "Prediction of Cyclic Stress–Strain Property of Steels by Crystal Plasticity Simulations and Machine Learning," Materials, **12**(22), p. 3668.

[44] Beniwal, A., Dadhich, R., and Alankar, A., 2019, "Deep Learning Based Predictive Modeling for Structure-Property Linkages," Materialia, **8**, pp. 100–435.

[45] Ali, U., Muhammad, W., Brahme, A., Skiba, O., and Inal, K., 2019, "Application of Artificial Neural Networks in Micromechanics for Polycrystalline Metals," Int. J. Plast., **120**, pp. 205–219.

[46] Proust, G., and Kalidindi, S. R., 2006, "Procedures for Construction of Anisotropic Elastic–Plastic Property Closures for Face-Centered Cubic Polycrystals Using First-Order Bounding Relations," J. Mech. Phys. Solids, **54**(8), pp. 1744–1762.

[47] Gel'fand, I. M., Minlos, R. A., and Shapiro, Z. Y., 2018, *Representations of the Rotation and Lorentz Groups and Their Applications*, Courier Dover Publications, Garden City, NY.

[48] Adams, B. L., Wright, S. I., and Kunze, K., 1993, "Orientation Imaging: The Emergence of a New Microscopy," Metall. Trans. A, **24**(4), pp. 819–831.

[49] Wu, X., Proust, G., Knezevic, M., and Kalidindi, S., 2007, "Elastic–Plastic Property Closures for Hexagonal Close-Packed Polycrystalline Metals Using First-Order Bounding Theories," Acta Mater., **55**(8), pp. 2729–2737.

[50] Fast, T., Knezevic, M., and Kalidindi, S. R., 2008, "Application of Microstructure Sensitive Design to Structural Components Produced From Hexagonal Polycrystalline Metals," Comput. Mater. Sci., **43**(2), pp. 374–383.

[51] Knezevic, M., Kalidindi, S. R., and Mishra, R. K., 2008, "Delineation of First-Order Closures for Plastic Properties Requiring Explicit Consideration of Strain Hardening and Crystallographic Texture Evolution," Int. J. Plast., **24**(2), pp. 327–342.

[52] Badrinarayanan, V., Kendall, A., and Cipolla, R., 2017, "Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," IEEE Trans. Pattern Anal. Mach. Intell., **39**(12), pp. 2481–2495.

[53] Du, C., and Gao, S., 2017, "Image Segmentation-Based Multi-Focus Image Fusion Through Multi-Scale Convolutional Neural Network," IEEE Access, **5**, pp. 15750–15761.

[54] Vardhana, M., Arunkumar, N., Lasrado, S., Abdulhay, E., and Ramirez-Gonzalez, G., 2018, "Convolutional Neural Network for Bio-Medical Image Segmentation With Hardware Acceleration," Cognit. Syst. Res., **50**(Special Issue:Deep Learning approaches for Cognitive Systems), pp. 10–14.

[55] Milletari, F., Navab, N., and Ahmadi, S.-A., 2016, "V-net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, Oct. 25–28, IEEE, pp. 565–571.

[56] Davies, A., Serjeant, S., and Bromley, J. M., 2019, "Using Convolutional Neural Networks to Identify Gravitational Lenses in Astronomical Images," Mon. Not. R. Astron. Soc., **487**(4), pp. 5263–5271.

[57] Jha, D., Singh, S., Al-Bahrani, R., Liao, W. K., Choudhary, A., De Graef, M., and Agrawal, A., 2018, "Extracting Grain Orientations From EBSD Patterns of Polycrystalline Materials Using Convolutional Neural Networks," Microsc. Microanal., **24**(5), pp. 497–502.

[58] Cang, R., Li, H., Yao, H., Jiao, Y., and Ren, Y., 2018, "Improving Direct Physical Properties Prediction of Heterogeneous Materials From Imaging Data Via Convolutional Neural Network and a Morphology-Aware Generative Model," Comput. Mater. Sci., **150**, pp. 212–221.

[59] Cecen, A., Dai, H., Yabansu, Y. C., Kalidindi, S. R., and Song, L., 2018, "Material Structure-Property Linkages Using Three-Dimensional Convolutional Neural Networks," Acta Mater., **146**, pp. 76–84.

[60] Dahl, G. E., Sainath, T. N., and Hinton, G. E., 2013, "Improving Deep Neural Networks for Lvcsr Using Rectified Linear Units and Dropout," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, Canada, May 26–31, IEEE, pp. 8609–8613.

[61] Ciregan, D., Meier, U., and Schmidhuber, J., 2012, "Multi-Column Deep Neural Networks for Image Classification," 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, June 16–21, IEEE, pp. 3642–3649.

[62] Feng Ning, D., Piano, F., Bottou, L., and Barbano, P. E., 2005, "Toward Automatic Phenotyping of Developing Embryos From Videos," IEEE Trans. Image Process., **14**(9), pp. 1360–1371.

[63] Hadsell, R., Sermanet, P., Ben, J., Erkan, A., Scoffier, M., Kavukcuoglu, K., Muller, U., and LeCun, Y., 2009, "Learning Long-Range Vision for Autonomous Off-Road Driving," J. Field Rob., **26**(2), pp. 120–144.

[64] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., and Kuksa, P., 2011, "Natural Language Processing (Almost) From Scratch," J. Mach. Learn. Res., **12**, pp. 2493–2537.

[65] Bottou, L., 2010, "Large-Scale Machine Learning With Stochastic Gradient Descent," *Proceedings of COMPSTAT'2010*, Y. Lechevallier, and G. Saporta, eds., Springer, Berlin, pp. 177–186.

[66] Groeber, M. A., and Jackson, M. A., 2014, "Dream. 3D: A Digital Representation Environment for the Analysis of Microstructure in 3D," Int. Mater. Manuf. Innov., **3**(1), p. 5.

[67] Smith, M., 2009, *ABAQUS/Standard User's Manual, Version 6.9.*, Dassault Systèmes Simulia Corp, Providence, RI.

[68] Kalidindi, S. R., Bhattacharyya, A., and Doherty, R. D., 2004, "Detailed Analyses of Grain–Scale Plastic Deformation in Columnar Polycrystalline Aluminium Using Orientation Image Mapping and Crystal Plasticity Models," Proc. R. Soc. Lond. Ser. A: Math. Phys. Eng. Sci., **460**(2047), pp. 1935–1956.

[69] Khan, A., Sohail, A., Zahoora, U., and Qureshi, A. S., 2020, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," Artif. Intell. Rev., **53**(8), pp. 5455–5516.

[70] Hecht-Nielsen, R., 1989, "Theory of the Backpropagation Neural Network," International 1989 Joint Conference on Neural Networks, Washington, DC, June 18–22, Vol. 1, pp. 593–605.

[71] Kingma, D. P., and Ba, J., 2015, "Adam: A Method for Stochastic Optimization," 3rd, International Conference on Learning Representations, Conference Track Proceedings, Y. Bengio and Y. LeCun, eds., San Diego, CA, May 7–9.