Extracting dislocation microstructures by deep learning

Yuqi Zhang and Alfonso H.W. Ngan[¶]

Department of Mechanical Engineering, University of Hong Kong, Pokfulam Road, Hong Kong

[¶] Corresponding author (email: hwngan@hku.hk)

Abstract

The microstructure of dislocations can be accessed by the total density of dislocations or the density of geometrically necessary dislocations (GND). The total dislocation density determines the flow strength of a crystal, which, in the case of high dislocation contents, is a quantity very difficult to measure accurately. On the other hand, related to crystal rotations, the GND densities are conveniently measured from electron diffraction experiments or calculated via simulations. Here, a novel and modern approach is proposed to understand the microstructures of dislocations based on deep learning, which estimates the total density of dislocations from a given density of GND distributions. In this method, the convolutional neural networks (ConvNets) are applied to extract the hidden information in the GND distribution maps to understand the microstructures of dislocations. It is demonstrated that the pre-trained ConvNets can be used to predict the distribution of total dislocation density from a small GND density map. Moreover, this technique is further developed to post-process real EBSD images for α-Fe to estimate the average total dislocation density, which corresponds to stress increments from a Taylor hardening assumption that is in good agreement with experimental values. Compared with previous methods involving much effort to track individual dislocations or other quantities, the present machine learning method is quick and convenient to use.

Keywords

Dislocation density; geometrically necessary dislocations; convolutional neural networks; machine learning

1. Introduction

Dislocations are the main carriers of plasticity in most crystalline materials. The well-known Taylor's formula relates the flow stress σ to the dislocation density ρ_{total} as $\sigma \propto \sqrt{\rho_{\text{total}}}$, and so ρ_{total} is one of the most important parameters in discussing crystal strength and plasticity. Traditionally, methods available for obtaining ρ_{total} include image-based methods such as transmission electron microscopy (TEM) or electron channeling contrast imaging (ECCI) (Dunlap et al., 2018; Gutierrez-Urrutia and Raabe, 2012), and diffraction-based methods such as X-ray, synchrotron X-ray and neutron diffraction. In addition to these methods, the density $\boldsymbol{\rho}_{\text{GND}}$ of geometrically necessary dislocations (GNDs) can be accurately determined by correlating with the lattice rotation and elastic strain obtainable by conventional electron backscattered diffraction (EBSD) (Calcagnotto et al., 2010; Dahlberg et al., 2014; Das et al., 2018; Demir et al., 2009; Hardin et al., 2013; Konijnenberg et al., 2015; Pantleon, 2008; Zhu et al., 2016) or high-resolution EBSD (HR-EBSD) (Chen et al., 2018; Jiang et al., 2015, 2013; T. J. Ruggles et al., 2016; T.J. Ruggles et al., 2016; Wilkinson et al., 2010). This EBSD-based approach is appealing not only because it captures the appropriate length scale for the heterogeneous deformation of metals, but also because the available mathematical formulas linking $\boldsymbol{\rho}_{\text{GND}}$ with the experimental data are mathematically stringent and trustworthy. However, the measured $\boldsymbol{\rho}_{\text{GND}}$ is not directly linked to the mechanical properties of the materials since, as mentioned above, the flow stress is defined by ρ_{total} rather than by $\boldsymbol{\rho}_{\text{GND}}$ according to Taylor's formula.

In fact, ρ_{total} and $\boldsymbol{\rho}_{\text{GND}}$ are geometrically and physically linked since they are the vector and scalar summation, respectively, of the density of the same group of individual dislocations (Ngan, 2017). Therefore, the total dislocation density distribution should be solvable based on the GND density map along with the appropriate boundary conditions. Recently, statistical models (Kalácska et al., 2017; T. J. Ruggles et al., 2016) have been proposed to estimate the total density based only on the local value of

$\boldsymbol{\rho}_{\mathrm{GND}}$, but then information hidden in the *distribution map* of $\boldsymbol{\rho}_{\mathrm{GND}}$ in the neighborhood is not exploited. Since $\boldsymbol{\rho}_{\mathrm{GND}}$ is based on individual dislocations which are lines connected in space, enhanced accuracy in the estimate of ρ_{total} should be achievable if the $\boldsymbol{\rho}_{\mathrm{GND}}$ distribution map in the neighborhood, rather than the local value of $\boldsymbol{\rho}_{\mathrm{GND}}$, is used.

For an image-like data structure (e.g., a $\boldsymbol{\rho}_{\mathrm{GND}}$ map), although information is stored locally in each pixel, it is the interconnection of the information in the neighboring pixels, rather than the data in a given pixel, that decides the meaning (such as the ρ_{total}) of the image. The advantage of ConvNets, as compared with other signal-processing techniques, lies in its characteristic structure (shared weights and multiple layers) that allows the essential information hidden in the neighboring pixels to be clearly identified. Due to this special advantage, ConvNets have recently been successfully employed in analyzing scanning electron microscopy (SEM) and EBSD images for microstructural recognition (Azimi et al., 2018; Chowdhury et al., 2016; DeCost et al., 2017; Hagita et al., 2018; Heinrich et al., 2017; Ling et al., 2017; Liu et al., 2016; Lubbers et al., 2017).

In this work, a new method is presented to understand the microstructure of dislocations, by estimating the total dislocation density ρ_{total} from the $\boldsymbol{\rho}_{\mathrm{GND}}$ map, with deep learning techniques (Lecun et al., 2015), especially the convolutional neural networks (ConvNets) (Krizhevsky et al., 2012; Szegedy et al., 2016). Furthermore, the trained ConvNets are applied to estimate the average total dislocation densities for real EBSD images from α-Fe samples. Compared with the previous methods to obtain total densities of dislocations, the present method does not require tracking individual dislocations or carrying out extra experiments or simulations.

2. Methodology

2.1 Problem statement

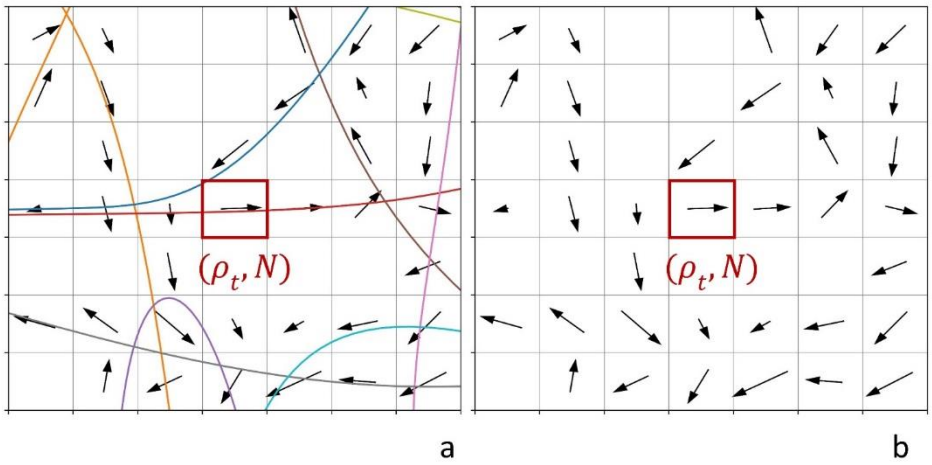

Fig. 1. (a) Random dislocation lines generated in a 7×7 matrix of pixels; arrows indicate the GND density vector in each pixel. (b) The GND density map of (a) used in the datasets with its corresponding (ρ_t, N).

A new framework for estimating the total dislocation density ρ_{total}, in terms of obtaining an explicit relation between ρ_{total} and $\boldsymbol{\rho}_{\text{GND}}$, is presented here. Based on the dislocation theory, the GND density and the total density are connected via

$$\begin{cases} \boldsymbol{\rho}_{\text{GND}}^{\alpha} = \sum_i \boldsymbol{\rho}_i^{\alpha} \\ \boldsymbol{\nabla} \cdot \boldsymbol{\rho}_i = 0 \\ \rho_{\text{total}}^{\alpha} = \sum_i |\boldsymbol{\rho}_i^{\alpha}| \end{cases} \tag{1}$$

where $\boldsymbol{\rho}$ is the density vector, α specifies the slip system, and i indicates the i^{th} individual dislocation in the slip system. Eqn. (1) suggests that both $\rho_{\text{total}}^{\alpha}$ and $\boldsymbol{\rho}_{\text{GND}}^{\alpha}$ depend on the characteristics of individual dislocations. Interestingly, when the coarse-graining resolution (the grid resolution in Fig. 1) is in nanometers, the distribution of $\boldsymbol{\rho}_{\text{GND}}^{\alpha}$ becomes an accurate topography of the individual dislocations, so

that the magnitude of $\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha}$ equals to $\rho_{\mathrm{total}}^{\alpha}$. Even though the coarse-graining length scale is in micrometers, information about individual dislocations can still be extrapolated from the GND distribution map, and we seek a general function *f* of the form

$$\rho_{\mathrm{total}}^{\alpha}(\mathbf{x}) = f(\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha}(\Omega)) \tag{2}$$

where **x** is the position on the slip plane at which the total density is estimated, and Ω is a neighborhood area around **x**. In other words, the purpose is to estimate the hidden information of $\rho_{\mathrm{total}}^{\alpha}$ from the $\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha}$ distribution regardless of the measurement resolution.

To achieve this goal, the model schematically shown in Fig. 2 is established. Given a large GND density distribution, $\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha}$, as the vector field shown in Fig 2(a), for each pixel, a small region of 7×7 pixels in its neighborhood is cropped as shown in the orange and green squares. The density $\rho_{\mathrm{total}}^{\alpha}$ in the center pixel is defined as $\rho_{\mathrm{total}}^{\alpha} = l/(s^2 b)$ where l is the dislocation line length in the pixel, s is the pixel length, and b is the slip plane thickness. The limited size of 7×7 pixels for Ω will introduce noise in the prediction but such noise can be reduced by additionally incorporating the number of dislocations $N_{\mathrm{total}}^{\alpha}$ inside each pixel as a model parameter, and hence the model parameters are specified as $(\rho_{\mathrm{total}}^{\alpha}, N_{\mathrm{total}}^{\alpha})$, which are hereafter written as (ρ_t, N) for short. Therefore, if the (ρ_t, N) for each pixel can be predicted, then the full (ρ_t, N) distribution map can be estimated.

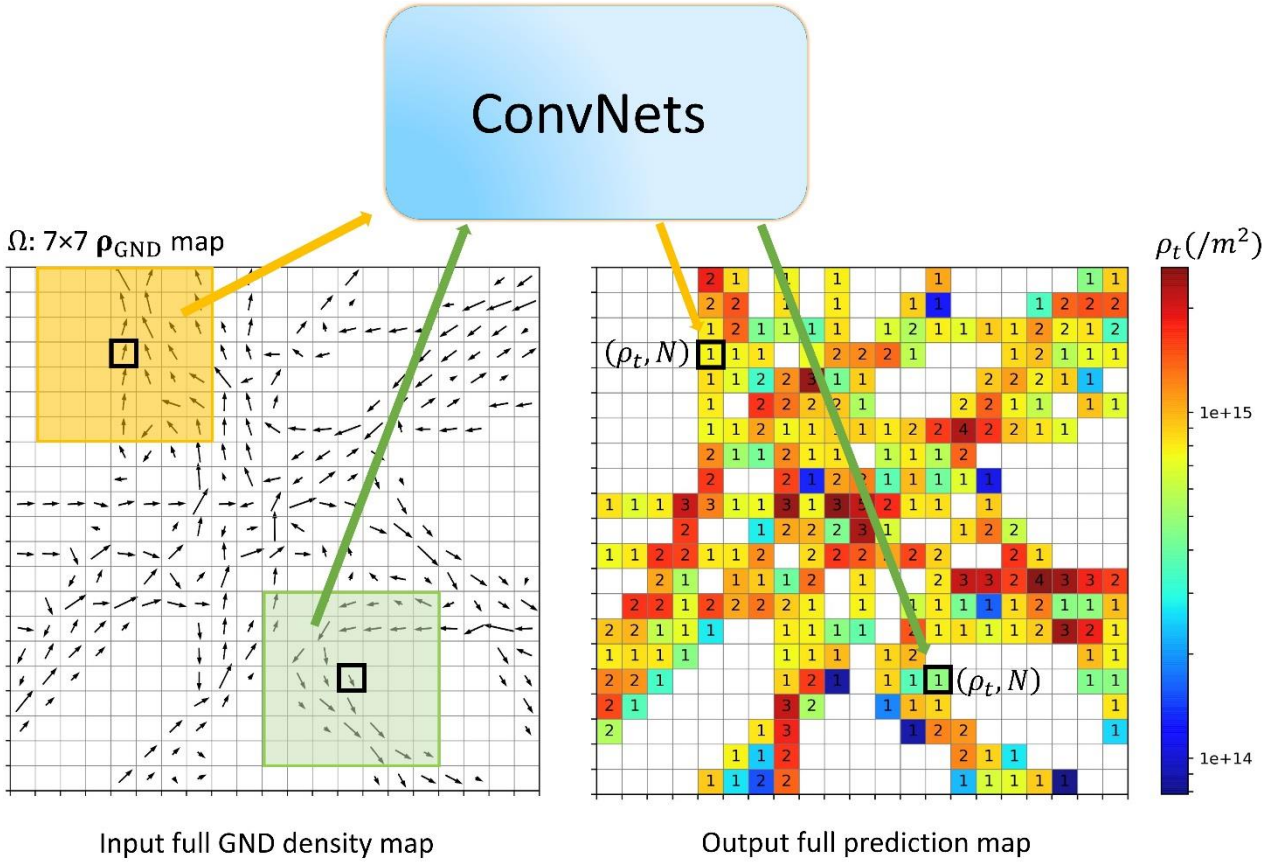

Fig. 2. Schematic of using a GND density map to predict the distribution of (ρ_t, N) via ConvNets where ρ_t is the total dislocation density and N the number of dislocations. The distribution of vector $\boldsymbol{\rho}_{\mathrm{GND}}$ shown as arrows is the input. For each pixel (outlined by black thick lines), a cropped 7×7 pixel grid (orange or green domain) serves as the area Ω used to predict (ρ_t, N). The full prediction of (ρ_t, N) in the right panel is the output. The color filling each pixel indicates the value of the ρ_t and the integers indicate the values of N.

2.2 Dataset generation

In the present machine learning approach, the input-output data are used to train, verify and test neural networks which are then used to make future predictions. To obtain such data for the slip plane specified

by α, dislocation lines are generated randomly as curved lines as shown in Fig. 2. To begin with, a 21×21 grid is generated and the middle point in each edge of the pixel is located as the approximate starting or finishing point of a dislocation passing through the pixel. Then, a set of random points from one boundary of the grid to another boundary is established. From the starting point to the ending point, each point looks for its next connecting point by choosing randomly from 5 neighbors except its predecessor point. B-spline is then interpolated through these points to represent a single dislocation line. A property of B-spline line is that the interpolated line does not necessarily pass through the given points, so the resulting dislocation line will be free from the initial grid. Finally, a random number of dislocation lines are generated in the 21×21 grid and a 7×7 cropped map is made as shown in Fig. 2. Then, the dislocation density vector $\boldsymbol{\rho}_i^{\alpha}$ for each line i is calculated, and the GND density vector in each pixel is obtained by vector summation, $\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha} = \sum_i \boldsymbol{\rho}_i^{\alpha}$, as shown in Fig. 1(b). The number of individual dislocations, N, and the total density ρ_t of each pixel are also obtained. We used a pixel size s of 3/7 μm and slip plane thickness b of 3 Å for calculating the individual dislocation line length l_i, and the density $\boldsymbol{\rho}_i(\mathbf{x}) = l_i \boldsymbol{\xi}_i / (s^2 b)$, where $\mathbf{x}$ is the position and $\boldsymbol{\xi}_i$ is the direction vector of the line i. With such a map constructed to generate data relating GND density $\boldsymbol{\rho}_{\mathrm{GND}}$ maps to the (ρ_t, N) of the center pixel, supervised learning is used to obtain the function in eqn. (2) for future predictions. Although the present method incorporates enough randomness, future studies may adopt discrete dislocation dynamics to generate datasets that may be more realistic.

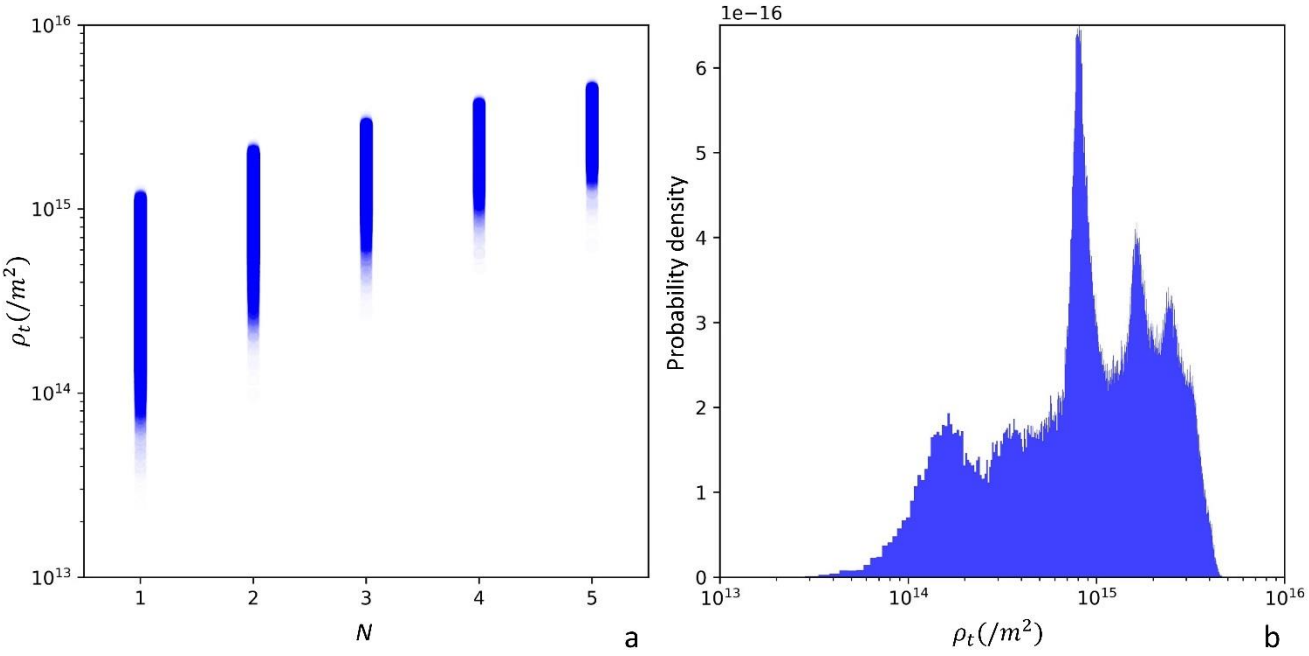

Fig. 3 (a) The total density of dislocation ρ_t vs. the number of dislocations N for the generated datasets. The discrete data points are each plotted with a transparency of 0.01, so that overlapping data points would show higher contrast, in order to indicate the distribution of the data. (b) The distribution of ρ_t for the datasets.

In a supervised learning sample, a pair of the 7×7 pixel GND density maps are input and its corresponding (ρ_t, N) in the center pixel as output. A GND density map is in the shape of 7×7×2, where the two channels are the edge and the screw components of the GND. Since the arrow in the GND map, as shown in Fig. 2(a), indicates the direction of the GND density, the edge and screw components can be either positive or negative while the value of the total density ρ_t cannot be negative. The whole dataset generated randomly consists of 5×10^5 samples covering the cases of N = 1, 2, 3, 4, 5 for training, $5\times2\times10^4$ samples for validation and $5\times2\times10^4$ samples for testing. Although in reality, the distribution of N may not be uniform amongst the samples, balanced datasets are used here to avoid biased predictions. Cases for

$N > 5$ are not considered since the real pixel size of the GND density map can always be reduced to exclude such cases with large N, and the purpose here is to show the working principles of this method, rather than aiming at a particular dislocation density value. From the generated dataset, ρ_t increases with the number of the dislocations inside the pixel N, and the Pearson correlation coefficient between N and ρ_t is 0.898. The two variables are highly correlated but certain deviations, as shown in Fig. 3(a), exist. The deviations are due to the finite size of the discrete pixels, so that the distribution and the shape of the dislocation lines inside the pixels are different. The distribution of ρ_t in Fig. 3(b) shows ρ_t in the range of 10^{13} to 10^{16} m^{-2}, with a mean value of 10^{15} m^{-2}.

2.3 Convolutional neural network

Classification (identifying which category a new observation belongs to) and regression (predicting a quantity for a new observation) are performed on the dataset generated above using the fully ConvNet, the architecture of which is summarized in Table A1 in appendix. The network contains 17 convolutional layers and 1 fully connected layer for 6 outputs with 343434 parameters in total. The loss function used here is a combination of classification loss (cal), softmax cross-entropy for N, and regression loss (reg) for ρ_t,

$$\begin{cases} \text{reg} = \dfrac{1}{m}\displaystyle\sum_{i=1}^{i=m}[\rho_t(i) - \widehat{\rho_t}(i)]^2 \text{ or} \\ \text{reg} = \dfrac{1}{m}\displaystyle\sum_{i=1}^{i=m}\{\log[\rho_t(i)+1] - \log[\widehat{\rho_t}(i)+1]\}^2 \\ \text{cal} = \dfrac{1}{m}\displaystyle\sum_{i=1}^{i=m}\sum_{j=1}^{j=5} p(i,j)\log[\hat{p}(i,j)] \\ \text{loss} = (1-\text{ratio})\cdot\text{reg} + \text{ratio}\cdot\text{cal} \end{cases} \quad (3)$$

where m is the number of the samples, p is the result after the softmax function, the ratio from 0 to 1 is used for the loss function, and the hat variables $\hat{p}$ and $\widehat{\rho_t}$ indicate the predictions from the networks. Two different regression losses are adopted in the training because the ρ_t distributes in a wide range as suggested in Fig. 3(b). The mean absolute percentage error (MAPE) of ρ_t,

$$\text{MAPE} = \frac{1}{m}\sum_{i=1}^{i=m}\frac{|\rho_t(i) - \widehat{\rho_t}(i)|}{\rho_t(i)} \tag{4}$$

and the classification accuracy of N are applied to measure the prediction accuracy of the ConvNets.

Before feeding the networks, the magnitude of the $\boldsymbol{\rho}_{\text{GND}}$ is added as the third channel along with the two existing channels of $\boldsymbol{\rho}_{\text{GND}}$ for edge and screw components. Each of the learnable layers are followed by the batch normalization transform (Ioffe and Szegedy, 2015), and the Relu (rectified linear unit) activation function expected for the output layer. Considering that the image size from the dataset is only 7×7 pixels, the valid padding type for the last six convolutional layers reduces the dimension gradually and the same padding type for previous convolutional layers is also adopted. A valid padding for a convolutional layer implies no padding, and will decrease the height and width of the map if the filter size > 1, while the same padding will pad zero around the map which can keep the height and width unchanged. A dropout rate of 0.5 is used for two dropout layers (Srivastava et al., 2014) in the training process. We apply the stochastic gradient descent and Adam optimizer with a training rate of 10^{-3}, and a mini-batch size of 64 in the training process. The validation dataset is used for hyperparameter tuning and early stopping to prevent overfitting. The convolutional neural network is implemented in the Tensorflow (Abadi et al., 2016) framework using the Python language. Python scripts for dataset generation, training ConvNets, test ConvNets, and trained model can be found at "https://github.com/Ninazizi/estimate_dislocation_density".

3. Training the ConvNets

Table 1. Models trained with different ratios for the loss function (see eqn. (3))

	Mean square error		Mean square logarithmic error	
ratio	MAPE of ρ_t	*N* Accuracy	MAPE of ρ_t	*N* Accuracy
0	44.3%	-	43.3%	-
0.25	40.7%	61.0%	40.8%	62.8%
0.5	41.9%	62.0%	**40.4%**	62.3%
0.75	42.2%	62.6%	41.2%	62.0%
1	-	**62.7%**	-	**62.7%**

The results are reported in Table 1 for two kinds of regression error, namely, the mean square error and the mean square logarithmic error, with different ratios applied to the loss function. The smaller the mean absolute percentage error (MAPE) of ρ_t and the larger the accuracy of *N*, the better the model. The major aim of the ConvNets is to predict ρ_t with the mean absolute percentage error as small as possible, and the smallest MAPE of ρ_t is achieved for the mean square logarithmic error in the model with the loss-function ratio of 0.5. Recalling from eqn. (3) that $\text{loss} = (1 - \text{ratio}) \cdot \text{reg} + \text{ratio} \cdot \text{cal}$, the models with a pure regression (ratio = 0) loss function achieve the worst MAPE of ρ_t among ratio of 0, 0.25, 0.5, 0.75, indicating that incorporating a classification (0 < ratio < 1) loss in the loss function helps generalize the model to predict ρ_t. The highest accuracy of *N* is achieved in the model with a loss function being exclusively the classification error (ratio = 1). The GND map alone does not uniquely indicate the ρ_t or *N* distribution when the pixel size is in the range of microns, since not all information is perfectly stored in the $\boldsymbol{\rho}_{\text{GND}}$ after coarse graining as described in eqn. (1), especially when a 7×7 pixel map size is used. This explains why in the present attempt, the MAPE of ρ_t does not converge to zero and the accuracy of *N* does not converge to 100%, and the ConvNets with 343,434 parameters trained with 500,000 samples still suffer from overfitting. To reduce the overfitting, joint (ρ_t, N) is applied in the loss function, and the

lowest MAPE of ρ_t occurs at the loss function ratio of 0.5 in Table 1. The generality of the ConvNets is enhanced by incorporating information for both ρ_t and N. Incorporating N in the training datasets and penalizing large deviations in the prediction of N in the training loss function (the classification loss in eqn. (3)) improve the generalized accuracy of ρ_t, as shown in Table 1. Involving the extra variables in the training datasets and using the combined loss function improve generalization, which is in the spirit of "multi-task learning" in machine learning.

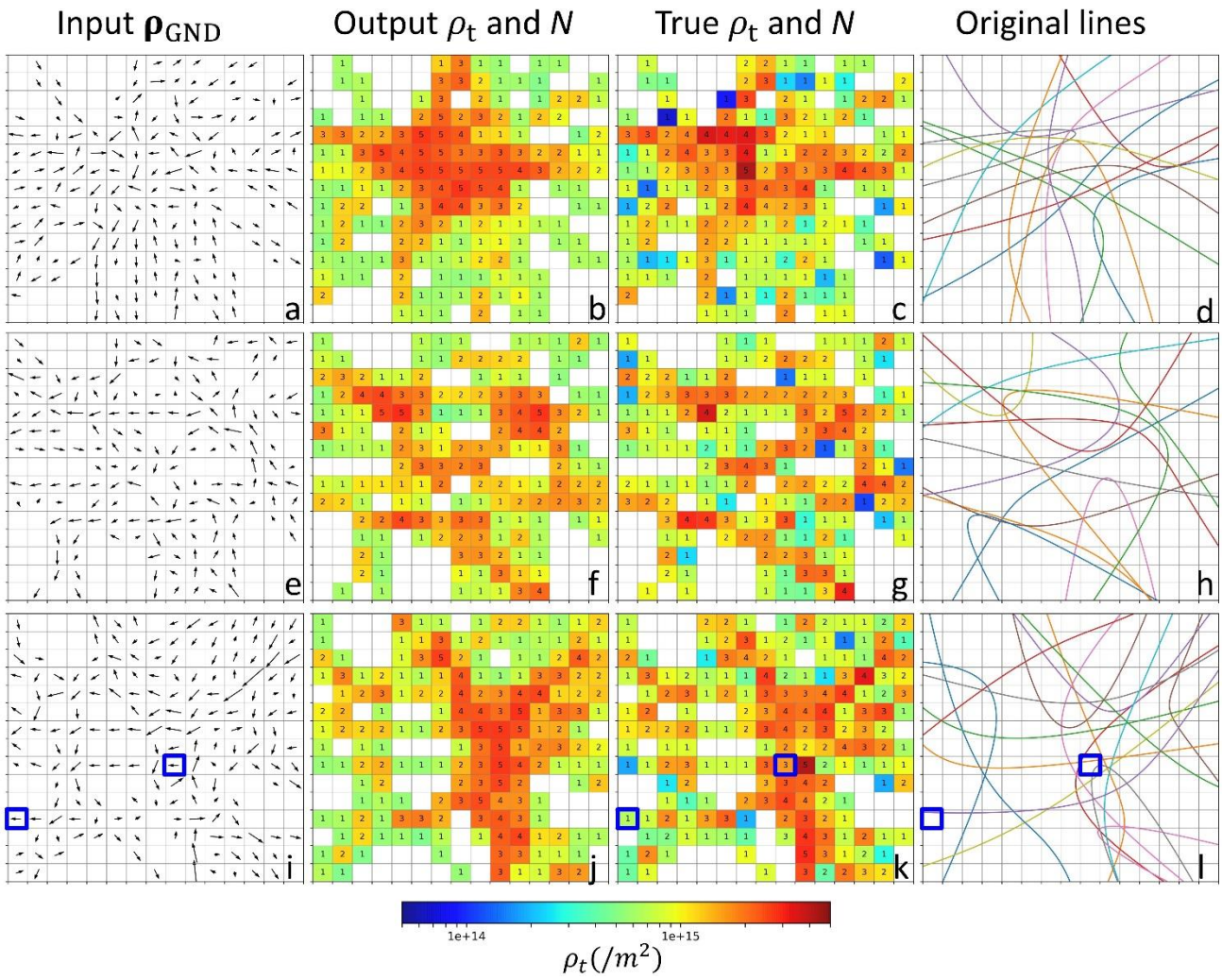

Fig 4. Examples of predicted total dislocation maps from the trained ConvNets. In the first column, (a), (e), (i) are the input of the $\boldsymbol{\rho}_{\mathrm{GND}}$ map. In the second column, (b), (f), (j) show the output (ρ_t, N) from the

trained ConvNets with the total dislocation density in color and the N in value. In the third and fourth columns, the figures show the true (ρ_t, N) values and true dislocation lines respectively. The blue boxes in (i) mark pixels with similar $\boldsymbol{\rho}_{\mathrm{GND}}$ but the corresponding boxes in (k) show quite different values of ρ_t and (l) shows the absolute different distributions of real lines.

Applying the ConvNets with the lowest mean square logarithmic error and loss-function ratio of 0.5, some prediction examples were obtained as shown in Fig. 4. Applying the trained ConvNets and the method shown in Fig. 2, the distribution of the total dislocation density can be estimated from the GND densities. The results in Fig. 4 show that the present method captures the distribution of ρ_t and provides an indicator for the complexity of the dislocation lines from the label N, i.e. the number of dislocations inside each pixel.

4. Application to real GND maps from EBSD experiments

Conventional EBSD measurements at spatial resolution (step size) of 0.1 μm were conducted for two α-Fe samples with different deformation strains (sample 1 with 4% and 10% strain, and sample 2 with 1%, 6%, and 10% strain). The GND densities are calculated from the EBSD measurements via the MTEX toolbox (Bachmann et al., 2010) based on methods proposed by Pantleon (Pantleon, 2008). The distributions of GND densities, $\sum_\alpha \left|\rho_{\mathrm{gnd}}^{\alpha}\right|$ where α indicates the slip system, are presented in Fig 5(b)-5(f). To apply the ConvNets for predicting the total density, the cropped 7×7 areas where the GND densities within some slip system satisfy the divergence-free requirement of $\boldsymbol{\nabla} \cdot \boldsymbol{\rho}^{\alpha} = 0$ were first selected. In practice, the numerical divergence for slip system α at each pixel is calculated by a finite difference method and if the sum of the absolute values of the divergences for all the 7×7 pixels in slip system α is

smaller than a critical value of 2×10^{11} m^{-2} /step size, the cropped 7×7 area is defined as satisfying the divergence-free requirement for slip system α. Then, ~800×7×7×2 GND densities maps are fed into the trained ConvNets and the average of ~800 prediction outputs is used as the average total dislocation density, ρ_t, for the specimen. Fig. 5(b-f) shows the cropped areas satisfying the divergence-free requirement in the GND map, with central pixels marked by cyan crosses. For each 200×200 pixel image shown in Fig 5(b-f), the number of cropped 7×7 images satisfying the requirement are ~800. Table 2 records the GND densities $\sum_\alpha|\rho^\alpha_{\text{gnd}}|$ and the predicted total densities ρ_t from the trained ConvNets.

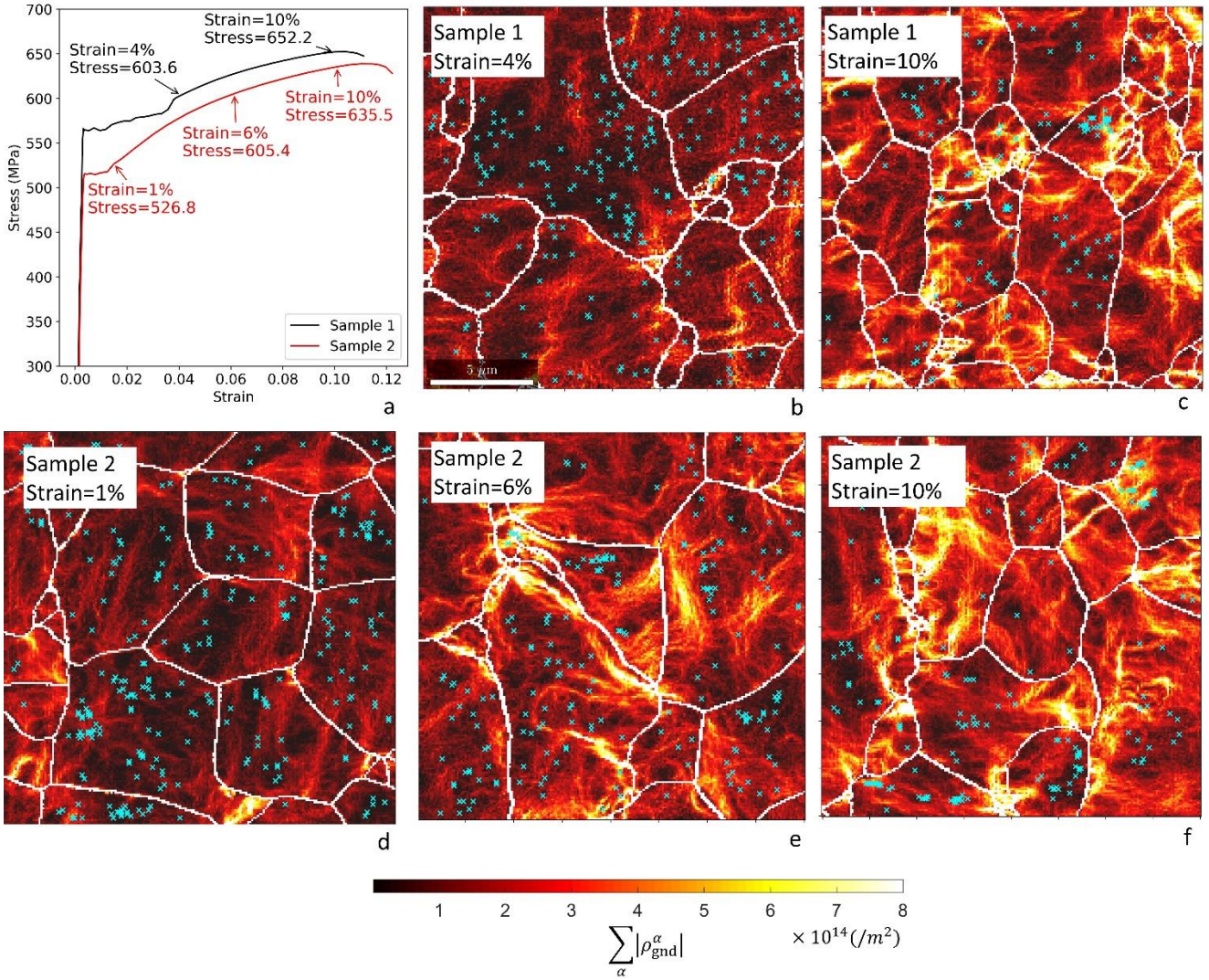

Fig. 5. (a) Stress-strain curves of two α-Fe samples. (b, c) GND densities for sample 1 at 4% and 10% strain respectively. (d, e, f) GND densities for sample 2 at 1%, 6% and 10% strain respectively. The cyan crosses in (b, c, d, e, f) indicate the places satisfying the divergence-free (continuity) requirement. The GND densities are calculated from all slip systems as $\sum_\alpha \left|\rho_{\mathrm{gnd}}^\alpha\right|$ where α indicates each slip system.

Table 2. GND densities, total densities, and experimental stresses for each sample.

	Sample 1		Sample 2		
Strain	4%	10%	1%	6%	10%
$\sum_\alpha \left|\rho_{\mathrm{gnd}}^\alpha\right|$ (/m^2)	1.81×10^{14}	2.85×10^{14}	1.50×10^{14}	2.48×10^{14}	2.98×10^{14}
ρ_t (/m^2)	9.28×10^{14}	1.16×10^{15}	8.80×10^{14}	1.17×10^{15}	1.26×10^{15}
Experimental stress (MPa)	603.6	652.2	526.8	605.4	635.5

To assess the predicted total dislocation densities, Taylor's formula for polycrystals is employed:

$$\sigma = \sigma^* + M\alpha'\mu b\sqrt{\rho} \tag{5}$$

where σ is the stress, σ^* is the critical stress, M is the Taylor factor, α' is a constant, μ is the shear modulus, σ is the burgers vector, ρ is the dislocation density. For each sample, a change $\Delta\sigma$ in the tensile stress from the stress-strain curve should be correlated with a change in the dislocation density as

$$\Delta\sigma = M\alpha'\mu b\Delta\left(\sqrt{\rho}\right) \tag{6}$$

Using the constants for α-Fe, namely, $\alpha' = 0.23$ (Kassner, 2004), b =2.482 Å, μ = 82 GPa, M = 2.86 (calculated from the EBSD data by MTEX (Bachmann et al., 2010)), in Fig. 6 the results of $\Delta\sigma$ calculated

from ρ_t and $\sum_\alpha|\rho_{\text{gnd}}^\alpha|$ are compared with the $\Delta\sigma$ from the tensile test experiments which are also recorded in Table 2. For all three situations, the predicted results from the current machine learning approach are better than the results from $\sum_\alpha|\rho_{\text{gnd}}^\alpha|$. In sample 1, the $\Delta\sigma$ calculated from ρ_t is closer to the true value of 48.6 MPa measured in the experiment, and both the values of $\Delta\sigma$ from $\sum_\alpha|\rho_{\text{gnd}}^\alpha|$ and from ρ_t are not very far from the experimental value. In sample 2 with larger stress increments, both results from $\sum_\alpha|\rho_{\text{gnd}}^\alpha|$ and ρ_t show underestimation of $\Delta\sigma$, but still the $\Delta\sigma$ from ρ_t is closer to the true value than that from $\sum_\alpha|\rho_{\text{gnd}}^\alpha|$. Moreover, when the stress increments calculated from the GND densities are close and fail to distinguish the hardening effects in the two situations (sample 1 from 4% to 10% strains, and sample 2 from 1% to 6% strains), the stress increments calculated from the estimated total densities clearly indicate that the latter situation has a stronger strain hardening effect than the first situation, which is in accordance with the tensile experiments. Since the ConvNets are never fed with the information of stress in the training process, its ability to predict the stress increment demonstrates the potential of the present machine learning approach.

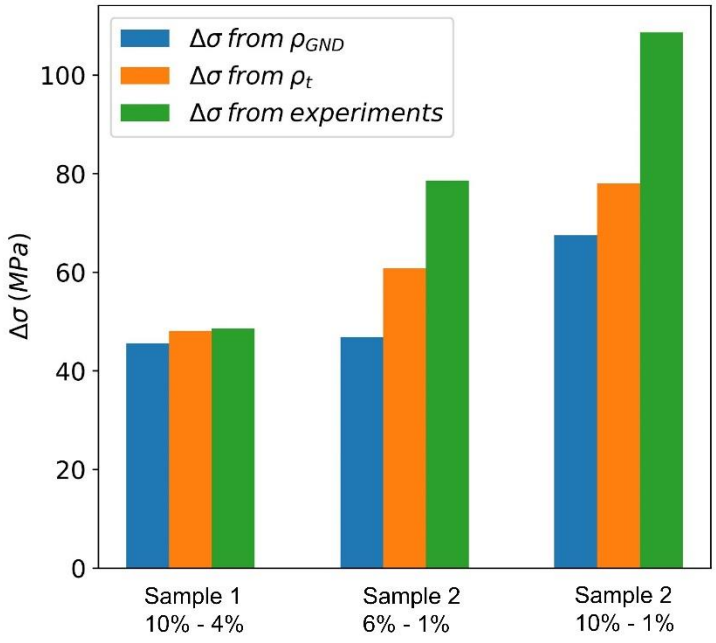

Fig. 6. Comparison among stresses increments ($\Delta\sigma$) calculated from GND density (ρ_{GND}) and total density of dislocations (ρ_t) by Taylor's law and the stress increments measured in experiments of three different situations: sample 1 between strain=10% and strain=4%, sample 2 between strain=6% and strain=1%, sample 2 between strain=10% and strain=1%.

5. Discussion

The present work first establishes a methodology to understand the microstructure of dislocations, namely, to estimate the total density of dislocation based on the distribution of the GND densities in the slip plane, and then the trained ConvNets are applied in post-processing the real EBSD data to estimate the total densities of dislocations. To establish the methodology in the first part, we focus on the essential property of dislocation lines, namely, that they must be continuous in the crystal. To apply the trained ConvNets to real EBSD data, the usable area is restricted according to the continuity requirement $\boldsymbol{\nabla} \cdot \boldsymbol{\rho}_i = 0$. Both attempts in the present study are around the core property of dislocation lines, i.e. the continuity requirement, which was neglected in previous studies (Kalácska et al., 2017; T. J. Ruggles et al., 2016). After fully digesting the hidden information from the continuity requirement using machine learning, a new approach is provided to understand the dislocation microstructures. The ConvNets is trained on the hypothetical 2D examples of dislocation lines on a single slip plane. In the application on the EBSD based GND maps, the divergence-free requirement which rejects terminating dislocations will select cropped small maps with slip planes close to the EBSD scanning plane. Since different crystals have different orientations, the application process can be seen as 2.5D.

Building on the continuity requirement of dislocation lines, the present methodology estimates the total density of a given point, $\rho^{\alpha}_{\mathrm{total}}(\mathbf{x})$, from the GND density in the neighborhood area (Ω) around that point,

$\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha}(\Omega)$, rather than only from the GND density at that point $\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha}(\mathbf{x})$. Utilizing the continuity property avoids establishing a locally one-to-one relationship between the total density and the GND density of dislocations (Kalácska et al., 2017; T. J. Ruggles et al., 2016). As Fig. 4(i, k, l) shows, two pixels with similar $\boldsymbol{\rho}_{\mathrm{GND}}$ may have very different values of ρ_t and configuration of dislocation lines. Therefore, eqn. (2), namely, $\rho_{\mathrm{total}}^{\alpha}(\mathbf{x}) = f\left(\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha}(\Omega)\right)$, serves to map the $\boldsymbol{\rho}_{\mathrm{GND}}$ in a certain area Ω around **x** to the ρ_t at position **x**, which is a more valid approach than the previous attempts (Kalácska et al., 2017; T. J. Ruggles et al., 2016) in which the total density is estimated from the value of $\boldsymbol{\rho}_{\mathrm{GND}}$ at the same location **x** only, i.e. $\rho_{\mathrm{total}}^{\alpha}(\mathbf{x}) = g\left(\boldsymbol{\rho}_{\mathrm{GND}}^{\alpha}(\mathbf{x})\right)$. By utilizing the information contained in the $\boldsymbol{\rho}_{\mathrm{GND}}$ map within a neighborhood Ω of the current pixel and the convolutional neural networks, the present method fully exploits the hidden geometrical information in the map, and higher accuracy can be achieved, as shown in the example of Fig. 4(j) which predicts distinct ρ_t values for pixels with similar $\boldsymbol{\rho}_{\mathrm{GND}}$ in Fig. 4(i).

The application of the present ConvNets to the real EBSD measurement in Fig. 5 is one step forward to illustrate the present approach. The ConvNets require the input cropped images to satisfy the divergence-free condition, which restricts the number of qualifying cropped maps. However, the predicted outputs of ρ_t reflect the change in strength of the samples. It should be noted that the step size in the EBSD measurement was 0.1 μm which is in the range of dislocation spacing $(1/\sqrt{\rho})$, which means that the total density is of the same magnitude as the GND density and so are the predicted results shown in Table 2. The selection of the step size influences the measured GND density as well as the total density. The application of ConvNets requires the $\boldsymbol{\rho}_{\mathrm{GND}}$ from the EBSD map to be in the same range as the $\boldsymbol{\rho}_{\mathrm{GND}}$ in the training datasets, which also means that the ρ_t from the EBSD map should be in the same range as the ρ_t in the training datasets. The ConvNets are based on the statistical description of the training datasets, so it is best to apply the model to EBSD maps with a similar range of $\boldsymbol{\rho}_{\mathrm{GND}}$ or ρ_t. The present model gives

a good estimation over a mild change in strength (sample 1), but it underestimates the ρ_t over a larger strength increment (sample 2).

The noticeable discrepancy between the stress increment ($\Delta\sigma$) from estimated ρ_t and that from experimental measurement (Fig. 6) shows that the present method underestimates $\rho_t$and the discrepancy increases with ρ_t. Four possible reasons might be responsible: (1) The ConvNets and the distribution of ρ_t in the training dataset (Fig. 3b) make the predicted results distribute around $10^{15}/m^2$, instead of extreme values that would be needed for large $\Delta\sigma$. (2) The EBSD measurement and the unindexed grain boundary areas will not have predicted ρ_t values while the grain boundary areas are likely locations for dislocation pile-ups. (3) The grain interiors are less well represented in high-strain samples than in low-strain samples. (4) Since the core concept of the present approach concerns the continuity of dislocations lines, cropped maps with terminated dislocation line segments would not satisfy the divergence requirement and the present method did not consider dislocation junctions with terminating ends. As a result, discrepancies in the stress increment exist between the model predictions and the experiments.

Although only single phase specimen is investigated in the present study, our approach here is also able to analyze multi-phase situations, and the total densities of dislocations for different phases can be separately estimated. As indicated in Fig. 5(b), the qualified points (satisfying the divergence-free requirement) are distributed in different grains, and the present method enables estimation of the distribution of the total densities in each grain, although the mean value of total densities for all qualified points is presented here. Recently, the GND density calculation from EBSD measurement has attracted much attention in investigating the microstructure of dislocations. With the present method, given the calculated GND density, the estimation of the mean or distribution of the total dislocation density for the sample is possible without additional experiments. The deformation mechanism of the material can be understood better given the distribution of the total densities of the dislocations.

6. Conclusion

A novel approach was presented to predict the total dislocation densities from the distribution of GND densities through the machine learning method of convolutional neural networks. The convolutional neural networks are trained and then applied to post-process EBSD data. Unlike previous approaches involving phenomenological descriptions to roughly estimate the total dislocation densities locally, machine learning is used to analyze the local total density from the GND density map over a small neighborhood and achieved good accuracy. Although only the GND densities from one slip plane in a two dimensional case are considered at the training process, the estimation of the mean dislocation density is in accordance with macroscopic experimental stress increments by Taylor's formula. With the rapid developments in computer vision and machine learning, our new approach is promising in extracting the total dislocation density from geometrically necessary dislocation densities based on EBSD measurement, and will help explore plastic deformation from an understanding of the dislocation microstructures.

Appendix

A1. The architecture of the convolutional neural networks

Table A1. The architecture of the convolutional neural networks. The layer names are composed of layer types (Input, Conv, Dropout, and so on) and the numbers (0, 1_1,4, and so on).

Layer name	Connected to	Padding type	(Filter size), # of filters	Output size
Input_0				7×7×3
Conv_1_1	Input_0	Same	(3, 3), 25	7×7×25
Conv_1_2	Conv_1_1		(1, 1), 25	7×7×25
Conv_2_1	Input		(1, 1), 10	7×7×10
Conv_2_2	Conv_2_1	Same	(3, 3), 25	7×7×25
Conv_3_1	Input		(1, 1), 25	7×7×25
Maxpooling_3_2	Conv_3_1	Same	(3, 3)	7×7×25
Concatenate_4	Conv_1_2, Conv_2_2, Maxpooling_3_2			7×7×75
Conv_5	Concatenate_4	Same	(3, 3), 30	7×7×30
Conv_6	Conv_5	Same	(3, 3), 35	7×7×35
Conv_7	Conv_6	Same	(3, 3), 40	7×7×40
Conv_8	Conv_7	Same	(3, 3), 45	7×7×45
Conv_9	Conv_8	Same	(3, 3), 50	7×7×50
Conv_10	Conv_9	Same	(3, 3), 55	7×7×55
Conv_11	Conv_10	Same	(3, 3), 60	7×7×60
Conv_12	Conv_11	Same	(3, 3), 65	7×7×65
Conv_13	Conv_12	Valid	(2, 2), 70	6×6×70
Conv_14	Conv_13	Valid	(2, 2), 75	5×5×75
Conv_15	Conv_14	Valid	(2, 2), 80	4×4×80
Conv_16	Conv_15	Valid	(2, 2), 85	3×3×85
Conv_17	Conv_16	Valid	(2, 2), 90	2×2×90
Conv_18	Conv_17	Valid	(2, 2), 95	1×1×95
Dropout_19	Conv_18			1×1×128
Conv_20	Dropout_19	Valid	(1, 1), 128	1×1×128
Dropout_21	Conv_20			1×1×128

Flatten_22	Dropout_21			128
Fully_connected_23	Flatten_22			6

A2. The influence of the area of the cropped map

It is natural to assume that a larger neighborhood area Ω for the prediction will result in higher accuracy. However, a large Ω may introduce other complications in reality. The hidden geometrical information in a GND density map that is useful in the estimation of ρ_t involves those dislocation lines that are connected between the pixels within the area Ω. In reality, a variety of dislocation mechanisms may induce terminated instead of continuous lines in the area, such as Frank-Read source, climb and cross slip segments out of the slip plane, and so on. Therefore, if Ω is too large to include many terminating dislocation segments, the accuracy will be reduced rather than enhanced. Another reason not to use a large Ω is that the boundary pixels of the given uncropped $\boldsymbol{\rho}_{\mathrm{GND}}$ map do not have enough neighbor pixels to form a cropped map of area Ω. A boundary pixel is therefore associated with a mapping area smaller than the present $\Omega = 7 \times 7$ pixels, and the missing area will be larger for a larger Ω. For the present $\Omega = 7 \times 7$ pixels, each grain (assuming square in shape) should have at least 7×7 pixels in order to result in a valid cropped map for the machine learning model. Although the results for $\Omega = 7 \times 7$ pixels are not perfect, the use of similar sizes is still suggested in the future.

Acknowledgement

This project is funded by the Kingboard Endowed Professorship in Materials Engineering at the University of Hong Kong. The authors thank Dr. Mingxin Huang and Dr. Shihui He for supplying the experimental samples.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., Brain, G., 2016. TensorFlow: A System for Large-Scale Machine Learning TensorFlow: A system for large-scale machine learning. 12th USENIX Symp. Oper. Syst. Des. Implement. (OSDI '16) 265–284. https://doi.org/10.1038/nn.3331

Azimi, S.M., Britz, D., Engstler, M., Fritz, M., Mücklich, F., 2018. Advanced steel microstructural classification by deep learning methods. Sci. Rep. 8, 1–14. https://doi.org/10.1038/s41598-018-20037-5

Bachmann, F., Hielscher, R., Schaeben, H., 2010. Texture Analysis with MTEX – Free and Open Source Software Toolbox. Solid State Phenom. 160, 63–68. https://doi.org/10.4028/www.scientific.net/SSP.160.63

Calcagnotto, M., Ponge, D., Demir, E., Raabe, D., 2010. Orientation gradients and geometrically necessary dislocations in ultrafine grained dual-phase steels studied by 2D and 3D EBSD. Mater. Sci. Eng. A 527, 2738–2746. https://doi.org/10.1016/j.msea.2010.01.004

Chen, B., Jiang, J., Dunne, F.P.E., 2018. Is stored energy density the primary meso-scale mechanistic driver for fatigue crack nucleation? Int. J. Plast. 101, 213–229.

https://doi.org/10.1016/J.IJPLAS.2017.11.005

Chowdhury, A., Kautz, E., Yener, B., Lewis, D., 2016. Image driven machine learning methods for microstructure recognition. Comput. Mater. Sci. 123, 176–187. https://doi.org/10.1016/j.commatsci.2016.05.034

Dahlberg, C.F.O., Saito, Y., Öztop, M.S., Kysar, J.W., 2014. Geometrically necessary dislocation density measurements associated with different angles of indentations. Int. J. Plast. 54, 81–95. https://doi.org/10.1016/J.IJPLAS.2013.08.008

Das, S., Hofmann, F., Tarleton, E., 2018. Consistent determination of geometrically necessary dislocation density from simulations and experiments. Int. J. Plast. https://doi.org/10.1016/j.ijplas.2018.05.001

DeCost, B.L., Francis, T., Holm, E.A., 2017. Exploring the microstructure manifold: Image texture representations applied to ultrahigh carbon steel microstructures. Acta Mater. 133, 30–40. https://doi.org/10.1016/j.actamat.2017.05.014

Demir, E., Raabe, D., Zaafarani, N., Zaefferer, S., 2009. Investigation of the indentation size effect through the measurement of the geometrically necessary dislocations beneath small indents of different depths using EBSD tomography. Acta Mater. 57, 559–569. https://doi.org/10.1016/j.actamat.2008.09.039

Dunlap, B.E., Ruggles, T.J., Fullwood, D.T., Jackson, B., Crimp, M.A., 2018. Comparison of dislocation characterization by electron channeling contrast imaging and cross-correlation electron backscattered diffraction. Ultramicroscopy 184, 125–133. https://doi.org/10.1016/j.ultramic.2017.08.017

Gutierrez-Urrutia, I., Raabe, D., 2012. Dislocation density measurement by electron channeling contrast imaging in a scanning electron microscope. Scr. Mater. 66, 343–346. https://doi.org/10.1016/j.scriptamat.2011.11.027

Hagita, K., Higuchi, T., Jinnai, H., 2018. Super-resolution for asymmetric resolution of FIB-SEM 3D imaging using AI with deep learning. Sci. Rep. 8, 1–8. https://doi.org/10.1038/s41598-018-24330-1

Hardin, T.J., Adams, B.L., Fullwood, D.T., Wagoner, R.H., Homer, E.R., 2013. Estimation of the full Nye's tensor and its gradients by micro-mechanical stereo-inference using EBSD dislocation microscopy. Int. J. Plast. 50, 146–157. https://doi.org/10.1016/J.IJPLAS.2013.04.006

Heinrich, L., Bogovic, J.A., Saalfeld, S., 2017. Deep learning for isotropic super-resolution from non-isotropic 3d electron microscopy. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 10434 LNCS, 135–143. https://doi.org/10.1007/978-3-319-66185-8_16

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv Prepr. arXiv1502.03167.

Jiang, J., Britton, T. Ben, Wilkinson, A.J., 2015. The orientation and strain dependence of dislocation structure evolution in monotonically deformed polycrystalline copper. Int. J. Plast. 69, 102–117. https://doi.org/10.1016/j.ijplas.2015.02.005

Jiang, J., Britton, T.B., Wilkinson, A.J., 2013. Measurement of geometrically necessary dislocation density with high resolution electron backscatter diffraction: Effects of detector binning and step size. Ultramicroscopy 125, 1–9. https://doi.org/10.1016/j.ultramic.2012.11.003

Kalácska, S., Groma, I., Borbély, A., Ispánovity, P.D., 2017. Comparison of the dislocation density

obtained by HR-EBSD and X-ray profile analysis. Appl. Phys. Lett. 110. https://doi.org/10.1063/1.4977569

Kassner, M.., 2004. Taylor hardening in five-power-law creep of metals and Class M alloys. Acta Mater. 52, 1–9. https://doi.org/10.1016/j.actamat.2003.08.019

Konijnenberg, P.J., Zaefferer, S., Raabe, D., 2015. Assessment of geometrically necessary dislocation levels derived by 3D EBSD. Acta Mater. 99, 402–414. https://doi.org/10.1016/j.actamat.2015.06.051

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks. Adv. Neural Inf. Process. Syst. 1–9. https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007

Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Ling, J., Hutchinson, M., Antono, E., DeCost, B., Holm, E.A., Meredig, B., 2017. Building data-driven models with microstructural images: Generalization and interpretability. Mater. Discov. 10, 19–28.

Liu, R., Agrawal, A., Liao, W.K., Choudhary, A., De Graef, M., 2016. Materials discovery: Understanding polycrystals from large-scale electron patterns. Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016 2261–2269. https://doi.org/10.1109/BigData.2016.7840857

Lubbers, N., Lookman, T., Barros, K., 2017. Inferring low-dimensional microstructure representations using convolutional neural networks. Phys. Rev. E 96, 52111.

Ngan, A.H.W., 2017. Dislocation-density kinematics: a simple evolution equation for dislocation density involving movement and tilting of dislocations. MRS Commun. 1–8.

https://doi.org/10.1557/mrc.2017.66

Pantleon, W., 2008. Resolving the geometrically necessary dislocation content by conventional electron backscattering diffraction. Scr. Mater. 58, 994–997. https://doi.org/10.1016/j.scriptamat.2008.01.050

Ruggles, T.J., Fullwood, D.T., Kysar, J.W., 2016. Resolving geometrically necessary dislocation density onto individual dislocation types using EBSD-based continuum dislocation microscopy. Int. J. Plast. 76, 231–243. https://doi.org/10.1016/J.IJPLAS.2015.08.005

Ruggles, T.J., Rampton, T.M., Khosravani, A., Fullwood, D.T., 2016. The effect of length scale on the determination of geometrically necessary dislocations via EBSD continuum dislocation microscopy. Ultramicroscopy 164, 1–10. https://doi.org/10.1016/j.ultramic.2016.03.003

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J. Mach. Learn. Res. 15, 1929–1958. https://doi.org/10.1214/12-AOS1000

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2818–2826.

Wilkinson, A.J., Clarke, E.E., Britton, T.B., Littlewood, P., Karamched, P.S., 2010. High-resolution electron backscatter diffraction: An emerging tool for studying local deformation. J. Strain Anal. Eng. Des. 45, 365–376. https://doi.org/10.1243/03093247JSA587

Zhu, C., Harrington, T., Livescu, V., Gray, G.T., Vecchio, K.S., 2016. Determination of geometrically necessary dislocations in large shear strain localization in aluminum. Acta Mater. 118, 383–394.

https://doi.org/10.1016/j.actamat.2016.07.051