

DOCTORAL THESIS

CARNEGIE MELLON UNIVERSITY

Applied Machine Learning to Predict Stress Hotspots in Materials

*Submitted in partial fulfillment of the requirements for
the degree of
Doctor of Philosophy*

in

Materials Science and Engineering

Ankita Mangal

M.S., Materials Science and Engineering, Carnegie Mellon University
B.Tech.(Hons.), Metallurgical and Materials Engineering, Indian Institute of
Technology, Kharagpur

Carnegie Mellon University
Pittsburgh, PA

June 2018

Ankita Mangal

Applied Machine Learning to Predict Stress Hotspots in Materials

© 2018

This work is licensed under a [Creative Commons](https://creativecommons.org/licenses/by-sa/4.0/)
“Attribution-ShareAlike 4.0 International” license.



ADVISOR:

Prof. Elizabeth A. Holm (MSE)

COMMITTEE:

Prof. Anthony D. Rollett (MSE)

Prof. Marc De Graef (MSE)

Prof. Leman Akoglu (Heinz College, CMU)

Dr. Bryce Meredig (Stanford University, Citrine Informatics)

LOCATION:

Pittsburgh, PA

DATE:

June 2018

ABSTRACT

This work focuses on integrating crystal plasticity based deformation models and machine learning techniques to gain data driven insights about the microstructural properties of polycrystalline metals. An inhomogeneous stress distribution in materials leads to the development of *stress hotspots* in polycrystalline metals under uniaxial tensile deformation. We simulate uniaxial tensile deformation in synthetic microstructures to get full field solutions for local micromechanical fields (stress and strain rates). After identifying stress hotspots by thresholding stress values, we characterize their neighborhoods using metrics that reflect the local crystallography, geometry, and connectivity. This data is used to create input feature vectors to train a random forest learning algorithm, which predicts the grains that become stress hotspots. We are able to achieve an area under the receiving operating characteristic curve (ROC-AUC) of 0.82 for hexagonal close packed and 0.74 for face centered cubic materials. Inspired by the recent advances in the deep learning field, we also explore using these techniques to automatically extract long range microstructural descriptors. The results show the power and the limitations of the machine learning approach applied to the polycrystalline grain networks.

PUBLICATIONS

The work presented in this dissertation is based on the following publications:

- A. Mangal and E. A. Holm. "Applied Machine Learning to Predict Stress Hotspots I: Face centered cubic materials." accepted in *International Journal of Plasticity*, 2018, preprint: <https://arxiv.org/abs/1711.00118>
- A. Mangal and E. A. Holm. "Applied Machine Learning to Predict Stress Hotspots II: Hexagonal close packed materials." under review in *International Journal of Plasticity*, 2018, preprint: <https://arxiv.org/abs/1804.05924>
- A. Mangal and E. A. Holm. "A comparative study of feature selection methods for stress hotspot classification in materials." *Integrating Materials and Manufacturing Innovation* (2018): 1-9., <https://doi.org/10.1007/s40192-018-0109-8>
- A. Mangal and E. A. Holm. "A dataset of synthetic face centered cubic 3D polycrystalline microstructures, grain-wise microstructural descriptors and grain averaged stress fields under uniaxial tensile deformation", accepted in *Data in brief*
- A. Mangal and E. A. Holm. "A dataset of synthetic hexagonal close packed 3D polycrystalline microstructures, grain-wise microstructural descriptors and grain averaged stress fields under uniaxial tensile deformation", under review in *Data in brief*

The dataset created for this dissertation has been published as:

- A. Mangal and E. A. Holm. "Synthetic FCC 3D polycrystalline microstructures with grain-wise microstructural descriptors and stress fields under uniaxial tensile deformation", *Mendeley Data*, <http://dx.doi.org/10.17632/ss75fdg5dg.1>
- A. Mangal and E. A. Holm. "Synthetic HCP 3D polycrystalline microstructures with grain-wise microstructural descriptors and stress fields under uniaxial tensile deformation : Part One", *Mendeley Data*, <http://dx.doi.org/10.17632/kt8hfg4t2p.1>
- A. Mangal and E. A. Holm. "Synthetic HCP 3D polycrystalline microstructures with grain-wise microstructural descriptors and stress fields under uniaxial tensile deformation : Part Two", *Mendeley Data*, <http://dx.doi.org/10.17632/nsfn6tw295.1>

“You miss 100% of the shots you never take”
Wayne Gretzky

ACKNOWLEDGEMENTS

First and foremost, I'd like to thank my family, who have been so supportive of me throughout graduate school. I am indebted to my parents, Mr. Rakesh K. Mangal and Mrs. Mansi Mangal, who have taught me the value of education; and my little brother Sanket, I could not have made it without your love and encouragement.

I have been extremely lucky to have a great advisor and the thesis committee chair, Prof. Elizabeth A. Holm, you are an inspiration and an awesome person! Thank you for encouraging me in all my endeavors and sending me to so many conferences. I have learnt a lot from you and would love to work with you again in the future.

I'd like to thank Prof. Anthony Rollett for his valuable inputs with the FFT simulations, and Prof. Marc De Graef for his computing resources. I would like to express deep gratitude towards my dissertation committee members Dr. Anthony Rollett (MSE,CMU), Dr. Marc De Graef (MSE,CMU), Dr. Leman Akoglu (Heinz College,CMU) and Dr. Bryce Meredig (Citrine Informatics) for your suggestions and feedback. You have helped me shape a solid dissertation.

I'd like to thank Prof. Vincent Sokalski for supporting me through the first 2.5 years of graduate school and teaching me how to conduct research. I am deeply grateful for everyone who helped me build this dissertation: Dr. Sudipto Mandal. Dr. Tugce Ozturk, and Dr. Vahid Tari, thank you for your help with crystal plasticity simulations, I could not have built the dataset without you. Dr. Brian DeCost and Toby Francis: it was really helpful exploring the applications of deep learning for my work, which appears in Chapter 7.

A kind thank you to all the members of Holm Research group - past and present- Dr. Brian DeCost, Dr. Ankit Gupta, Dr. Philip Goins, Dr. Jonathan Humberson, Andrew Kitahara, Ian Chesser, Toby Francis and Anna Smith, it has been nice sharing the graduate student life with you, I had a great time working from and exploring different cafes! I want to give a big shout out to the MSE staff members : Jeanna, Kelly, Suzy, Marygrace and Dave for keeping things together and not letting us worry about any logistics.

I want to mention my undergraduate advisors Prof. Shiv Brat Singh (IIT Kharagpur) and Dr. Saurabh Kundu (Tata Steel), without whose encouragement, I would have never thought of pursuing a PhD. Their trust in my capabilities made me consider it.

To my friends at CMU: Dr. Pratiti Mandal, Dr. Arka P. Roy, Dr. Sudipto Mandal, Dr. Deepoo Kumar, Dr. Prabhat KC, Shivram Kashyap, Aditi Bhattacharya, Dr. Tugce Ozturk, Sudhanshu Nahata, Dr. Samikshya Subedi, Dr. Vignesh Sundar and everyone else: I feel lucky to know you guys and it has been awesome pursuing extracurricular activities and attending conferences together. We had so much fun in Salt Lake City, San Diego, Phoenix and of course in Pittsburgh! A special thank you to my good friends Tiyasa Mitra, Adhithi Shwetha Adhi, Dr. Manaal Faruqui and Nisarga Markandaiah, for hanging out, going on trips and giving me a strong moral support.

The following open source softwares were used in this work: scikit-learn.org, keras.io, tensorflow.org. The authors would like to acknowledge the computational facilities of the Materials Characterization Facility at CMU under grant MCF-677785. Ricardo Lebensohn of the Los Alamos National Laboratory is acknowledged for the use of the MASSIF (EVPFFT) code.

This work was funded by the United States National Science Foundation under grant number DMR-1307138 and DMR-1507830.

Contents

Abstract	ii
Publications	iii
Acknowledgements	iv
Contents	vi
List of Figures	ix
List of Tables	xiii
Abbreviations	xiv
1 Introduction	1
2 Crystal Plasticity Fundamentals	5
2.1 Crystallography	5
2.1.1 Miller Indices	6
2.1.2 Miller-Bravais Indices	6
2.2 Fundamentals of Crystal Plasticity	7
2.2.1 Slip Systems	7
2.2.2 Critically Resolved Shear Stress	8
2.2.3 Polycrystal Plasticity	10
2.2.3.1 Single Crystal Yield Surface	10
2.3 Crystal Plasticity Constitutive Models	11
2.3.1 Mean Field Crystal Plasticity Models	12
2.3.2 Image based Full Field Crystal Plasticity Models	13
2.3.2.1 Crystal Plasticity simulations using FFT	14
2.3.2.2 Elasto-Viscoplastic Fast Fourier Transform (EVPFFT)	15
2.3.2.3 Viscoplastic Fast Fourier Transform (VPFFT)	17
2.3.3 Summary	18
3 Machine Learning Fundamentals	19
3.1 Machine Learning algorithms	20
3.1.1 Decision Tree Models	20
3.1.2 Random Forest Tree Models	21
3.1.2.1 Variable Importance	23
3.1.3 Deep learning	23

3.1.3.1	The Perceptron	24
3.2	Model performance metrics	25
3.2.1	Accuracy, Precision and Recall	26
3.2.2	Area Under Receiving Operator Characteristic Curve (AUC)	27
3.3	Model Generalization	28
4	Predicting Stress Hotspots in Face Centered Cubic Materials	29
4.1	Introduction	29
4.2	Methods	31
4.2.1	Dataset Generation	31
4.2.1.1	Synthetic Microstructures	31
4.2.1.2	Simulating Uniaxial Tensile Deformation	32
4.2.1.3	Problem Formulation: defining stress hotspots	32
4.2.1.4	Relation Between Stress Hotspot locations	33
4.2.1.5	Effect of Microstructure evolution on Stress hotspots	33
4.2.1.6	Effect of Texture and Microstructure	34
4.2.2	Developing Microstructural Descriptors	35
4.2.2.1	Crystallographic Descriptors	35
4.2.2.2	Geometrical Descriptors: Grains and their neighborhood	38
4.2.3	Machine Learning Methods	39
4.2.3.1	Model Performance Metrics	39
4.2.3.2	Estimation of Model Generalization Error	40
4.2.3.3	Random Forest Model	41
4.2.3.4	Feature Importance Metrics	42
4.3	Results and Discussion	42
4.4	Conclusions	46
4.4.1	Contributions	47
5	Understanding Stress hotspots in Hexagonal Close Packed Materials	48
5.1	Introduction	48
5.2	Methods	50
5.2.1	Dataset Generation	50
5.2.2	Microstructural Descriptors	52
5.2.3	Machine Learning Methods	53
5.3	Results and Discussion	54
5.3.1	Role of competing slip systems in stress hotspot formation	58
5.4	Conclusions	61
5.4.1	Contributions	61
6	Feature Selection Techniques	63
6.1	Introduction	63
6.2	Methods	65
6.2.1	Dataset Studied	66
6.2.2	Feature Selection Methods	67
6.2.2.1	Filter Methods	67
6.2.2.2	Embedded Methods	68
6.2.2.3	Wrapper Methods	69

6.3	Results and Discussion	70
6.4	Conclusions	75
7	Predicting Material Failure Using Deep Learning	76
7.1	Introduction	76
7.1.1	Previous Work and Challenges	78
7.2	Dataset	79
7.2.0.1	Data Augmentation	80
7.3	Model Development and Results	81
7.3.1	3D neural network architectures	82
7.3.2	2D neural network architectures	85
7.4	Summary	87
8	Future work	89
A	Constitutive Parameters: Face Centered Cubic materials	92
B	Constitutive Parameters: Hexagonal Close Packed materials	95
C	Understanding Hotspot Neighborhoods	98
C.1	Transgranular network	98
C.2	Comparing Grain Neighborhoods	100
C.2.1	Effect of grain neighborhood on stress hotspots	102
C.3	Summary	106
D	Characterizing the Transgranular Network	107
D.1	Summary	112
	Bibliography	113

List of Figures

2.1	3 common atomic arrangements in crystals: a) BCC unit cell b) FCC unit cell and c) HCP unit cell. The unit cell for each case is outlined in solid black lines. d) Schematic showing Miller indices of planes and directions in a cubic crystal	6
2.2	Schematic showing a) Miller-Bravais indices for planes in HCP crystals: The green planes are related by 6-fold symmetry along the c axis, and belong to the family of $\{10\bar{1}0\}$ planes. b) Miller-Bravais indices for the basis vectors in HCP crystals are shown in red. The corresponding Miller indices of the basis vectors is shown in black.	7
2.3	Schematic for different slip systems in the 3 crystal structures: BCC, FCC and HCP. The slip planes are shaded and the slip directions are shown in red arrows.	8
2.4	A single crystal under uniaxial tension, highlighting the resolved shear stress τ on an arbitrary slip system	9
2.5	Representative volume element: The black dots are the Fourier points overlaid on an input microstructure image. Note that the actual grid is much finer than shown here.	14
3.1	Schematic of a decision tree model	21
3.2	Schematic of the random forest algorithm.	22
3.3	a) Schematic of a single layer perceptron b) Different activation functions in a perceptron	25
3.4	Schematic of an artificial neural network constructed from perceptron units	26
3.5	Schematic of 3.5a: the ROC curve and AUC metric. The diagonal line represents random prediction, 3.5b: learning curve for the high bias case and 3.5c: learning curve for the high variance case	27
4.1	Cross sections of a microstructure showing (a) Von Mises Stress field and (b) Stress hotspots obtained by thresholding on grain averaged stress field. The red regions correspond to regions of high stress or stress hotspots.	33
4.2	Representative figure showing the Von Mises stress vs. location in the microstructure at each deformation step. The bottom (blue) curve is at the lowest strain (0.01%) and the top (blue) curve is at the maximum strain (4%). The red arrows indicate how location of stress hotspots is not changing	34

4.3	Crystallographic descriptors used as features during machine learning. (a) Schematic of an Inverse pole figure for a cubic crystal. The distance of the [001] sample direction from the three crystal directions [001] (d2), [101] (d3) and [111] (d1) in the standard stereographic triangle are used as a feature. (b) Schematic of average misorientation ω_{avg} for the reference (light blue) grain with respect to its N contiguous neighbors (numbered 1 to 6) in a 2-D microstructure. ω_i is the misorientation between the reference grain and its i^{th} neighbor grain. (c) Slip Transmission Factor (m-prime, m') is a function of the dot product of the slip plane normals (incoming normal n_{in} and outgoing normal n_{out}), and the dot product of slip directions (incoming d_{in} and outgoing d_{out}). Adapted from [1]	37
4.4	(a) Representative pole figures for six different FCC textures. (b) Histograms of grain averaged stress in with different textures of FCC materials and (c) Learning curve for random forest model built on all textures of FCC materials	43
4.5	ROC curves of the mixed microstructure model for validation microstructures in each representative FCC texture. The AUC is the area under this curve.	43
4.6	Variable importance in FCC materials: Both texture and geometry derived microstructural features are selected by FeaLect algorithm. The importance scale is arbitrary and set by Fealect	44
5.1	Schematic of the different slip systems in a hexagonal close packed structure: basal {0001}[1120], prismatic {1010}[112] and pyramidal $\langle c+a \rangle$. When the tensile axis lies in the (10 $\bar{1}$ 1) pyramidal plane, the Schmid factor of prismatic $\langle a \rangle$ slip is higher than the basal $\langle a \rangle$ slip.	50
5.2	Representative textures for 8 different HCP textures, the corresponding scale bars show the texture intensity in MRD	51
5.3	Histograms of grain averaged stress with different textures in HCP materials with (a) Equal CRSS ratio and (b) Unequal CRSS ratio. The corresponding learning curves for Mixed-Micro model in HCP materials with (c) Equal CRSS ratio and (d) Unequal CRSS ratio	55
5.4	ROC curves of the mixed microstructure model for validation microstructures in each representative HCP texture for both the CRSS ratio cases. The AUC is the area under this curve.	56
5.5	FeaLect variable importance in HCP materials showing selected features for Equal CRSS (green) and Unequal CRSS (blue)	57
5.6	Cross section of a randomly textured 3-D equiaxed microstructure showing the Von Mises stress distribution under different SCYS topology regimes for a microstructure with random texture.	58
5.7	Correlation matrix for slip activities and Schmid factors	59
5.8	Pole figures of the initial random texture, hot grains in Equal CRSS case and hot grains in Unequal CRSS case. Note that different scale bars.	60
6.1	Pearson Correlation matrix between the target (EqVonMisesStress) and all the features	72
7.1	A typical CNN architecture showing an input image, 2 sets of convolutional, nonlinearity and pooling layers and a fully connected multilayer perceptron at the end to classify the input image into 5 classes.	78

7.2	Data Augmentation Pipeline	80
7.3	Maxpooling and Undo operation	83
7.4	Problems with 3D architectures. a) Checkerboard like artifacts in 3D-FCN outputs b) Vanishing gradient problems in 3D-CNN-hypercolumns and c) Regression output	83
7.5	a) Neural Net architecture for 3D Resnet_Grain-wise model	84
7.6	a) Neural network architecture for a 2D Residual network inspired pixel-wise segmentation model with 13 convolutional layers. b) Ground truth and predicted classes for validation data c) Confusion matrix for validation data. The total number of true labels in each bin are noted at the right, and total predicted labels per bin are noted at the top.	87
A.1	VPSC simulation fit to the experimentally observed stress-strain curve for OFHC Copper	94
A.2	Histogram of the important features distinguishing hot and normal grains in FCC	94
B.1	HCP α -Titanium fit	96
B.2	VPSC simulation fit to the experimentally observed stress-strain curve for alpha-Titanium.	96
B.3	Histogram of the important features distinguishing hot and normal grains	97
C.1	A 2-d microstructure. The grain boundary network is shown in black. It's dual, the transgranular network (shown in green) is constructed with grains as the nodes (drawn at grain centroids), and connections across each grain boundary. An ego graph of radius 1 (only first nearest neighbors) is highlighted in red.	99
C.2	Pole figure representing the texture of the microstructure chosen for analysis using spectral features. The tensile direction is Z (out of plane)	101
C.3	Density power law. Plots show total count of nodes vs. total count of edges in the egonet for all nodes (in log-log scales). The green line is least squares fit on the median values (black stars). Hot grains have more clique like neighborhoods.	101
C.4	Weight power law. Plots show total weight vs. total count of edges in the egonet for all nodes (in log-log scales). The green line is least squares fit on the median values (black stars). The neighborhood graphs for hot grains deviate from equation C.2	102
C.5	Eigenvector power law. Plots show the principal eigenvalue vs. total sum of weights in the egonet for all nodes (in log-log scales). The green line is least squares fit on the median values (black stars).	103
C.6	Histogram of the neighbor probability based features distinguishing hot and normal grains for 80 different kinds of microstructures	104
C.7	Average absolute Pearson correlation coefficient averaged over 80 different microstructures. The features derived from second nearest neighbors stack together at the top.	105
D.1	Cliques containing stress hotspots in a trangranular network form a percolating network.	108
D.2	Cliques containing stress hotspots in a transgranular network form a single connected component.	109

D.3	The largest connected component in the transgranular network for different stress threshold values. The red nodes are stress hotspots	111
D.4	Nodes below the threshold stress are deleted and the graph structure was studied. The critical point for connectivity percolation is shown by a vertical line.	112

List of Tables

4.1	Cross validation AUCs for FCC materials for mixed and partition models	42
4.2	Pearson Correlation Coefficients between features and Stress hotspots for FCC materials.	44
5.1	Cross validation AUCs (%) for mixed and partition models in equal and Unequal CRSS ratio case of HCP materials	54
5.2	Pearson Correlation Coefficients between features and Stress hotspots for HCP materials	55
6.1	Variable Importance Measures using different methods for HCP materials with Unequal CRSS. The gray shaded cells denote the features selected by the corresponding technique. The features describing grain geometry are shaded in green.	71
7.1	Summary of convolutional neural net development efforts	86
A.1	Single crystal elastic stiffness constants (in GPa) for Copper	92
A.2	Voce Hardening law parameters imitating Copper	93
A.3	Feature name descriptions	93
B.1	Voce Hardening law parameters for α -Titanium	96
B.2	Single crystal elastic stiffness constants (in GPa) for α -Titanium	96
C.1	Mean probability based feature values for n^{th} nearest neighbors	105
D.1	Connectivity percolation thresholds for site and bond percolation in different lattices and dimensions [2]	110

Abbreviations

AUC	A rea U nder C urve
BOR	B urgers O rientation R elationship
CFS	C orrelation-based F eature S election
CPFEM	C rystal P lasticity F inite E lement M ethod
CPFFT	C rystal P lasticity F ast F ourier T ransform
CRSS	C ritically R esolved S hear S tress
EVFFT	E lasto V isco P lastic F ast F ourier T ransform
FC	F ull C onstraints
FCC	F ace C entered C ubic
FEM	F inite E lement M ethod
FFT	F ast F ourier T ransform
HCP	H exagonal C lose P acked
HEDM	H igh E nergy X - R ay D iffraction M icroscopy
ICME	I ntegrated C omputational M aterials E ngineering
IPF	I nverse P ole F igure
MDF	M isorientation D istribution F unction
MGI	M aterials G enome I nitiative
ML	M achine L earning
RF	R andom F orest
RFE	R ecursive F eature E limination
ROC	R eceiving O perator C haracteristic
RVE	R epresentative V olume E lement
ODF	O rientation D istribution F unction
OFHC	O xygen F ree H igh T hermal C onductivity
VPFFT	V isco P lastic F ast F ourier T ransform
VPSC	V isco P lastic S elf C onsistent

Chapter 1

Introduction

Computational methods for developing microstructure-property relationships constitute an important part of materials and process design for high performance applications. These methods incorporate knowledge about material physics, experimental calibration and validation to help establish quantitative relationships between a material microstructure and the physical properties. In 2011, the Materials Genome initiative (MGI) was established to develop materials at an accelerated rate with the help of collaborations between material scientists and computer scientists [3]. The initiative aims to create a materials innovation infrastructure which integrates experimental tools, computational tools and digital data, to accelerate the materials development timeline from discovery to deployment. This has actualized the field of materials informatics, fusing data driven machine learning techniques with materials development. As a result, several data infrastructure platforms have come into existence [4], [5],[6]. Since then, there has been a flurry of research on computational materials discovery to search for stable compounds across the composition space [7], [8], [9], [10], [11]. However, understanding structure-property relationships by integrating machine learning methods with experimental and computational tools is an emerging area of research.

Metals and their alloys are widely used as structural materials. These materials require high toughness and ductility to maintain their structural integrity to meet engineering standards. Ductile fracture is one of the most common modes of failure in such materials, and has been widely researched [12], [13]. An important mechanism

for ductile fracture in metals and their alloys is by the nucleation, growth and coalescence of microscopic voids [14], [15]. The voids have a tendency to nucleate near defects and discontinuities in the structure, such as second phase particles in a matrix, near larger particles [16], twin boundaries, grain boundaries, vacancy clusters and slip bands. It has been shown that void nucleation is induced by stress, and voids can grow under stress by accumulating vacancies [17], [14]. Ductile inter-granular and faceted failures can also occur due to void nucleation resulting from the intersection of slip bands with grain boundaries and other slip bands in Oxygen Free High thermal Conductivity (OFHC) Copper [18], β Ti alloys [19] and α Ti-Al alloys. Ultimately, a complex interplay of microstructural features determines damage nucleation in materials. The ability to predict where voids could potentially form in a material during deformation, given its microstructure, can be leveraged to develop new processing paths to design materials more resistant to damage.

An applied stress on a polycrystalline material is heterogeneously distributed [20] between the grains, creating regions of stress accumulations, so-called stress hotspots. These regions are determined by the local characteristics in a microstructure. Rollett et. al. [21] have shown that stress hotspots in face centered cubic, randomly oriented equiaxed microstructures under uniaxial tension tend to form near the microstructural features, such as grain boundaries, triple and quadruple junctions and usually form in textures corresponding to maxima in Taylor factor as a function of orientation for that loading condition.

To understand the parameters behind inhomogeneous distribution of applied stress, we propose using statistical data mining methods to gain empirical relations between local microstructural characteristics and the stress hotspots. However, building a large dataset for statistical data mining through experiments is tedious and expensive. Simulations reduce the cost and development time of new materials by circumventing repetitive and expensive mechanical testing. Starting with the 3D microstructure of a material, we want to predict where the material is most probable to fail, and process it in a way to avoid the features causing these problems. The structure of a material changes a lot by the time a fracture starts to form in it. Simulating deformation in materials gives us the advantage of preserving both the initial and final structures, thus enabling us to “turn back the clock” and correlate hot spots to initial features.

Since void nucleation and stress hotspots are local events, it is important to choose a deformation model which will give full field solutions of stress and strain fields at each discretized point in a microstructure. We use the Elasto-Viscoplastic Fast Fourier Transform based (EVPFFT) crystal plasticity model [22] to simulate deformation in synthetic polycrystalline microstructures. The motivation is to find underlying microstructural characteristics that determine the formation of stress hotspots and use them to predict probable failure sites in similar materials. We use statistical descriptors of microstructures describing the crystallography (orientation distribution function, Taylor factor, Schmidt factor, n-point correlations), geometry (grain shape, grain boundary types) and connectivity (transgranular network, deep learning) which have not been used earlier to understand their correlation with hotspot locations.

The primary contributions of the present work is to explore new models for understanding structure-property relationships by developing microstructural descriptors and using them to build statistical learning models which predict a physical phenomenon. The models used here draw heavily from the field of machine learning and computer vision. It should be noted here that the models are developed using synthetic data, hence we cannot expect accurate predictions about experimental data from them. Instead, the work is dedicated towards developing an approach for generating microstructural descriptors to understand structure-property relationships.

This dissertation is organized into related, self contained chapters, some of which have been submitted for publication at the time of writing this thesis. Chapter 2 goes over the fundamentals of crystal plasticity, required to understand the dataset generation process and the physics of plastic deformation. In Chapter 3, I give a broad overview of machine learning techniques used in this work. Chapter 4 establishes and applies the framework of stress hotspot prediction for face centered cubic materials. This chapter shows how plastic deformation of ductile single phase polycrystals can be studied using a data science based approach. Chapter 5 extends this approach to hexagonal close packed (HCP) materials, which have a more complex deformation mechanism, and shows the merits of using a data driven approach. Chapter 6 is dedicated to understanding the feature selection techniques for getting data driven insights. Then we dive into applying deep learning methods for stress hotspot predictions in Chapter 7. Finally, in Chapter 8, I discuss the open question of extracting multiscale microstructural descriptors from a

deep learning framework, and attempt at identifying the future research directions and goals stemming from this dissertation.

Chapter 2

Crystal Plasticity Fundamentals

In this dissertation, we study plastic deformation of single-phase metals to understand where stress hotspots form. A computational model based on fundamentals of crystal plasticity [20] is used to generate full field discrete solutions for stress and strain rate micromechanical fields that develop inside polycrystalline microstructures. By using artificially generated microstructures as an input, a dataset of hotspots can be built for different materials and loading conditions. This chapter discusses the fundamentals of crystal plasticity and the constitutive models which were used to build the dataset.

2.1 Crystallography

Crystallography is the study of the spatial arrangement of atoms in crystalline solids. Crystalline solids have a periodic arrangement of atoms in space known as the crystal structure. The periodicity of this arrangement allows us to define a smallest possible repeating unit, the **unit cell**. The atoms or molecules are regularly arranged in three dimensional space on a **crystal lattice**. The position of any vector \mathbf{r} within the lattice can be defined as:

$$\mathbf{r} = u\mathbf{a} + v\mathbf{b} + w\mathbf{c} \quad (2.1)$$

where \mathbf{a} , \mathbf{b} and \mathbf{c} refer to the crystal reference frame or the crystal axes/ basis vectors and depend on the type of crystal under consideration. A combination of all possible angle and directions of basis vectors in 3-D leads to seven crystal axes systems and the

additional consideration of atomic arrangements leads to fourteen **Bravais Lattices** [23]. The three most common crystal structures are face-centered cubic (FCC), body-centered cubic (BCC) and hexagonal-close packed (HCP). The unit cells demonstrating the atomic arrangement of these structures are illustrated in figure 2.1, where the spheres represent atoms. The atoms in the structures are colored according to the sheet of spheres they correspond to in the 3-D tiling.

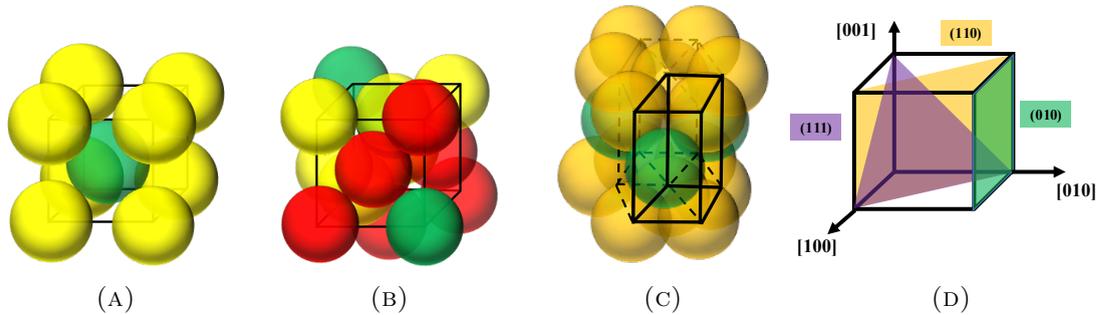


FIGURE 2.1: 3 common atomic arrangements in crystals: a) BCC unit cell b) FCC unit cell and c) HCP unit cell. The unit cell for each case is outlined in solid black lines. d) Schematic showing Miller indices of planes and directions in a cubic crystal

2.1.1 Miller Indices

Miller indices are used to specify the planes and directions in a crystal. The Miller indices of directions in a crystal are denoted on the basis of lattice vectors in square brackets $[hkl]$, where h , k and l are integers. The notation $\langle hkl \rangle$ is used to denote all directions related by symmetry. The Miller indices of a plane is defined using the reciprocal of the plane's intercept on the three crystal axes and are given as triplet of three integer values (hkl) . The family of symmetry related planes is denoted by $\{hkl\}$. Figure 2.1d shows (110) , (010) and (111) planes, and $[100]$, $[010]$ and $[001]$ directions in a cubic crystal.

2.1.2 Miller-Bravais Indices

In HCP materials, the crystallographic axes of the unit cell are not orthogonal. If Miller indices are used to denote the crystal planes, the crystallographically equivalent (symmetry related) planes have dissimilar Miller indices. By using four basis vectors $(a_1, a_2, a_3$ and $c)$, three of which are coplanar ($(a_1, a_2$ and $a_3)$); the Miller-Bravais indices for HCP planes can be derived such that symmetry related planes can be obtained via

permutations of the indices. These indices are written as $(hkil)$. Figure 2.2a shows the new basis vectors and the Miller-Bravais indices for the crystallographic planes, with the symmetry related planes highlighted in the same color. The Miller-Bravais indices for directions are expressed as $[uvw]$. Figure 2.2b shows the Miller indices and Miller-Bravais indices for the basis vectors in HCP crystals.

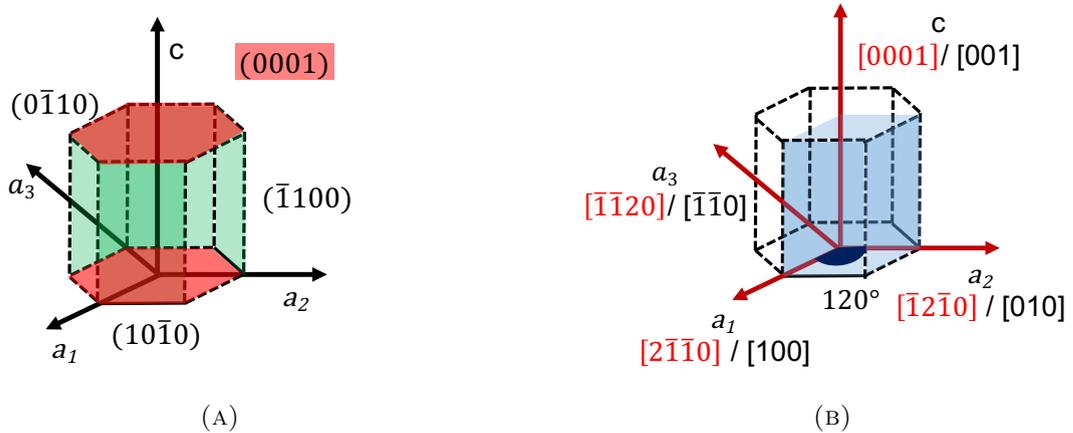


FIGURE 2.2: Schematic showing a) Miller-Bravais indices for planes in HCP crystals: The green planes are related by 6-fold symmetry along the c axis, and belong to the family of $\{10\bar{1}0\}$ planes. b) Miller-Bravais indices for the basis vectors in HCP crystals are shown in red. The corresponding Miller indices of the basis vectors is shown in black.

2.2 Fundamentals of Crystal Plasticity

2.2.1 Slip Systems

Plastic deformation in materials occurs due to the movement of a large number of dislocations. This movement causes the atoms to displace along a plane of one part of the crystal relative to other under a shear stress. This phenomenon is known as slip in materials. The shear stress required for dislocation motion to move the crystal by a step is proportional to that distance i.e. the Burger's vector. Hence, a small distance between atoms is favorable for slip. Thus slip is preferred in close packed planes and the dislocation movement occurs along the closest packed direction in that plane. This configuration is known as the **slip system** of a crystal structure. For example, in FCC materials the planes of $\{111\}$ family are closest packed and the $\langle 1\bar{1}0 \rangle$ crystallographic directions are the slip directions. In BCC materials, the slip system is $\{110\} \langle 111 \rangle$.

HCP materials have multiple slip systems on the prismatic, basal and pyramidal planes. Figure 2.3 demonstrates the different slip systems in these materials.

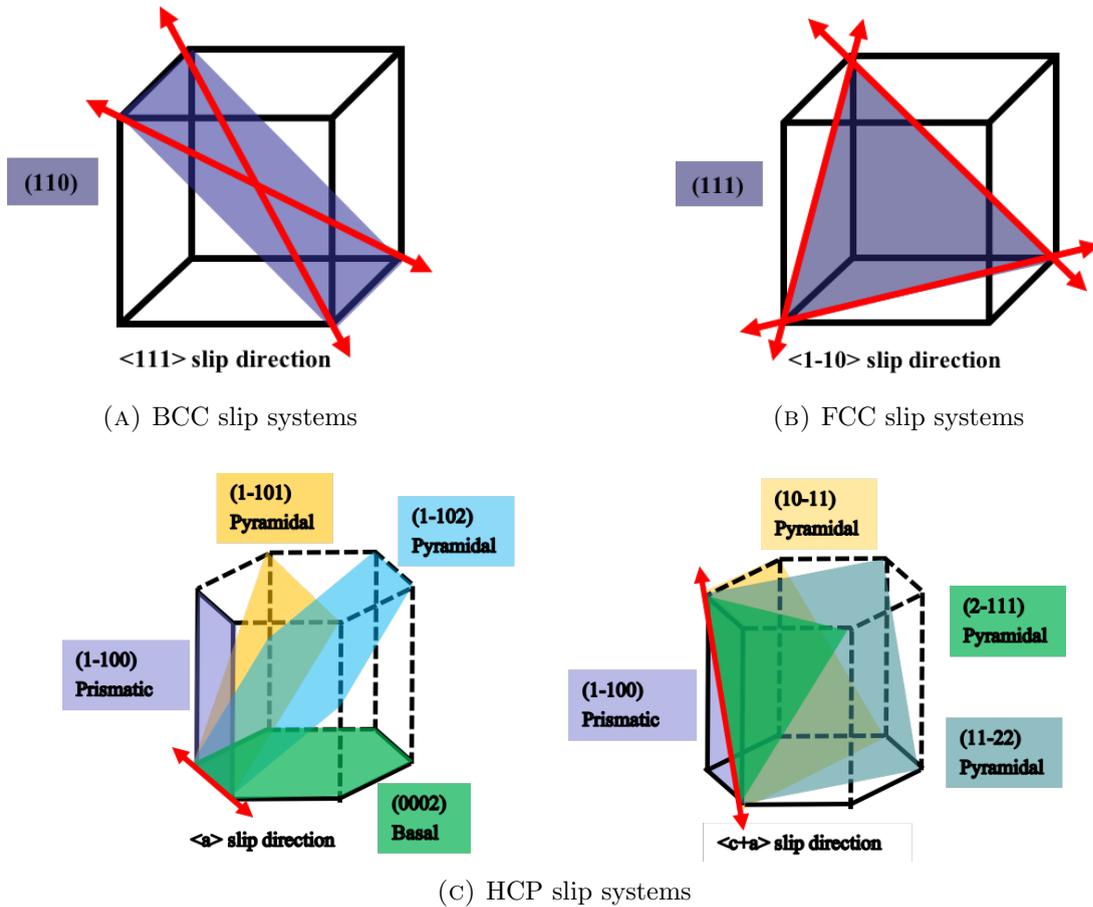


FIGURE 2.3: Schematic for different slip systems in the 3 crystal structures: BCC, FCC and HCP. The slip planes are shaded and the slip directions are shown in red arrows.

2.2.2 Critically Resolved Shear Stress

When a stress is applied to a material, the resolved shear stress in the crystal on the slip plane in the slip direction determines whether a given slip system will be active. For the case of uniaxial tensile deformation in a single crystal, the stress tensor is given by:

$$\boldsymbol{\sigma} = \sigma \hat{\boldsymbol{p}} \otimes \hat{\boldsymbol{p}} \quad (2.2)$$

where $\boldsymbol{\sigma}$ is the stress tensor, $\hat{\boldsymbol{p}}$ is the unit vector along the uniaxial tensile loading direction, and σ is the stress magnitude. Since plastic deformation occurs on the slip

systems, let us consider a slip system with slip plane normal unit vector \hat{n} and slip direction unit vector \hat{s} . As can be seen from figure 2.4, the resolved shear stress on the slip system is given by: $\tau = \hat{n} \cdot \boldsymbol{\sigma} \cdot \hat{s} = \sigma(\hat{p} \cdot \hat{n})(\hat{s} \cdot \hat{p})$ We can see that: $(\hat{p} \cdot \hat{n}) = \cos \phi$

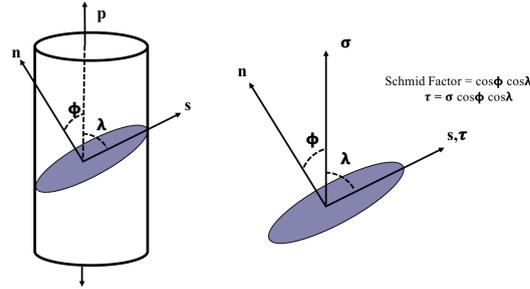


FIGURE 2.4: A single crystal under uniaxial tension, highlighting the resolved shear stress τ on an arbitrary slip system

and $(\hat{s} \cdot \hat{p}) = \cos \lambda$. The resolved shear stress on the slip system can be written as the **Schmid Law**:

$$\tau = \sigma \cos \phi \cos \lambda \quad (2.3)$$

The Schmid factor is thus defined as the cosine product, $\cos \phi \cos \lambda$, which is a measure of the optimal orientation of a slip system for deformation. Slip systems with higher Schmid factor have higher resolved shear stress for the same applied tensile stress. Plastic deformation occurs on the slip system where the resolved shear stress first exceeds the critical value required for plastic deformation. This critically resolved shear stress (CRSS), (τ_c) is a property of the crystal. As the material deforms, the CRSS increases and a hardening law describes the evolution of τ_c with strain.

Once a slip system becomes active, the crystal rotates such that the slip direction aligns with the loading axis and the corresponding Schmid factor becomes zero. When multiple slip systems with different CRSS values are present, for example in HCP materials with 3 deformation modes: prismatic slip, basal slip and pyramidal slip, a plastic anisotropy is introduced. The slip modes with a lower CRSS are favored and determine the crystal rotation. For HCP materials, there is an additional geometric anisotropy due to the hexagonal crystal structure.

2.2.3 Polycrystal Plasticity

Common materials used for engineering applications have a polycrystalline structure. These polycrystals consist of “grains”, which are regions similar to a single crystal. A grain is a three dimensional region in a polycrystal having the same phase, and the same crystal orientations within a tolerance. Grain boundaries are two-dimensional regions where two grains meet. For polycrystals, the deformation mechanism is more complex. In addition to the multiple slip systems present in a grain, polycrystal deformation is constrained by the neighborhood. These complex interactions lead to the development of “texture” in polycrystals under deformation. The texture describes the orientation distribution of individual crystals in a polycrystalline material [24].

2.2.3.1 Single Crystal Yield Surface

To understand the polycrystal plasticity and texture development in terms of single crystals, the concept of single crystal yield surface was developed. The stress and strain increment are symmetric 3x3 tensors, hence only consist of 6 independent components (σ_{ij}). During plastic deformation, the volume is conserved. Since hydrostatic stress (σ_{kk}) causes volume change, the plastic work becomes independent of it, and is only a function of the deviatoric stresses (σ'_{ij}), of which only 5 components are independent.

$$\sigma'_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{kk}\delta_{ij} \quad (2.4)$$

The Schmid law (equation 2.3) can be written in the 5 independent stress component space as:

$$\tau^s = |m_{ij}^s \sigma_{ij}| \quad (2.5)$$

where m_{ij}^s is the Schmid factor for a given slip system s for the stress component σ_{ij} . Thus, the Schmid law in the 5 dimensional stress space (2.5) defines a hyperplane, at a distance τ^s from origin. During plastic deformation, to accommodate a plastic strain increment with p components ($p \leq 5$), p slip systems need to be activated i.e. the resolved shear stress should be greater than or equal to the CRSS on p slip systems.

Thus the minimum required constraints for plastic deformation to occur is given by:

$$|m_{ij}^s \sigma_{ij}| \begin{cases} = \tau_c^s & \text{in } p \text{ independent slip systems} \\ \leq \tau_c^s & \text{in all others} \end{cases} \quad (2.6)$$

The **single crystal yield surface** (SCYS) is defined by equation 2.6 as an inner envelope of all the facets corresponding to the slip systems. The SCYS determines the shears that are activated in a grain in a polycrystal and depends on the CRSS ratios between deformation modes, as well as the stress state. Even if the CRSS of a mode is very high, it might be activated to complete the yield surface to achieve 5 independent slip modes. The SCYS has been analyzed and derived in detail for BCC materials in [25], for FCC materials in [26] and HCP materials in [27]. It has been shown [28] that the SCYS is topologically invariant in certain domains of CRSS ratios, and leads to a simplified analysis of deformation when slip modes harden at different rates. The CRSS ratios is defined with respect to the basal slip resolved shear strength (τ_{basal}) as:

$$CRSSRatio = Prismatic : Basal : Pyramidal = \frac{\tau_{prismatic}}{\tau_{basal}} : 1 : \frac{\tau_{pyramidal}}{\tau_{basal}} \quad (2.7)$$

2.3 Crystal Plasticity Constitutive Models

With the understanding of polycrystal plasticity, various models have been developed to predict the stress-strain response and the texture development in crystalline solids during deformation. The Sachs model [29] for polycrystalline deformation, developed in 1928, assumes iso-stress conditions, i.e. all grains are subject to the same stress tensor; and the activation of a single slip system. This model satisfies the stress equilibrium between grains at the grain boundaries, but not the strain compatibility conditions, and hence provides a lower bound prediction to macroscopic deformation of polycrystals. The Sachs model also has the defect that gaps and overlaps develop immediately as strain accumulates.

In contrast, the Taylor model [30] assumes iso-strain conditions, the simultaneous activation of multiple slip systems, and the same hardening rate for all slip systems in the microstructure. The strain compatibility condition at grain boundaries is satisfied by the existence of at least 5 active slip systems that minimize the plastic work. This full

constraints (FC) model provides an upper bound on achieving a required deformation under an applied load. The yield stress σ and plastic strain increment $\delta\epsilon$ of a polycrystal is determined on the basis of a suitable average of shear stress and strains in the individual grains. The Taylor factor M links the macroscopic applied stress with the individual shears in the grains as:

$$\sigma = M\tau \quad (2.8)$$

$$\delta\Gamma = M\delta\epsilon \quad (2.9)$$

where τ is the CRSS of the slip system and $\delta\Gamma$ is the sum of individual shear contributions

The Bishop Hill criterion in multiple stress states gives us a set of solutions (vertices) that can be activated to accommodate the given stress. Note that the Bishop Hill model applies only to cubic materials.

To find the slip rates on each slip system, instead of making an arbitrary choice of the vertex that satisfies the external slip, a rate sensitive slip equation is solved. This approach typically tries to “round the corners” of the single crystal yield surface. The stress state is chosen based on the imposed strain rate tensor.

$$\dot{\epsilon}^p(x) = \sum_{s=1}^N m^s(x) \dot{\gamma}^s(x) = \dot{\gamma}_0 \sum_{s=1}^N m_{ij}^s \left(\frac{m_{kl}^s : \sigma'_{kl}}{\tau^s} \right)^n \text{sgn}(m_{kl}^s : \sigma'_{kl}) \quad (2.10)$$

The slip (shear) rate on the individual slip system s can then be calculated while avoiding any assumptions to resolve ambiguity of which slip systems are activated.

$$\dot{\gamma}^{(s)} = \dot{\epsilon}_0 \left| \frac{m^{(s)} : \sigma^c}{\tau^{(s)}} \right| \text{sgn} \left(m^{(s)} : \sigma^c \right) \quad (2.11)$$

2.3.1 Mean Field Crystal Plasticity Models

To improve the agreement between experimental and calculated textures, a relaxed constraints Taylor model was developed by assuming that certain shear strain components can be neglected as they generate very small displacements in volumes, for example, in high aspect ratio grains, the shear strain in the short directions can be ignored. Esheby [31] derived the mean field homogenization problem in the form of an ellipsoidal inclusion in an infinite matrix for elastic materials.

The self-consistent models started with Hutchinson [32] in 1970 and were also developed by Molinari and Canova [33] in 1987. This approach was then extended to develop the viscoplastic self-consistent (VPSC) model at the Los Alamos National Laboratory for the simulation of large plastic deformation in polycrystalline aggregates [33], [34], [35] in 1993-1994. In this model, each grain is assumed to be an ellipsoidal inclusion in a viscoplastic homogeneous equivalent medium, which represents the average environment of that grain. The ellipsoidal grain behaves as a single crystal, deforming via crystallographic slip and the slip rate is a function of the critically resolved shear stress (CRSS) of that system. Thus the inhomogeneous microstructure is replaced via homogenization with a macroscopically isotropic medium. These models generally overestimate the texture [36]. Also, stress hotspots are a local phenomenon in the microstructure and the VPSC model neglects the effect of local fields. To understand the microstructure-property relationships better, a model that gives a full field solution to deformation is used.

2.3.2 Image based Full Field Crystal Plasticity Models

With better computing resources, it has become possible to solve for micromechanical fields that develop in a polycrystal. Both short-range and long-range grain interactions are considered and the micromechanical fields are resolved on a discrete grid. Crystal plasticity finite element models (CPFEM) were first developed in 1982 [37] and have been extensively used for predicting the texture and microstructure evolution of polycrystals. FE methods generally scale as $O(N^2)$, where N is the number of degrees of freedom in the system. The construction of meshes that conform to the grain boundaries in complex microstructures is difficult and time consuming. This difficulty in meshing, along with the large number of degrees of freedom required by CPFEM limits the use for complex problems to a representative volume element (RVE) consisting of few hundred grains.

In contrast, fast Fourier transform (FFT) based techniques circumvent the problems arising due to meshing in FEM simulations. FFT allows direct usage of measured or synthetic microstructure images due to the use of a regular grid. These algorithms are computationally efficient and scale as $O(N \log N)$, where N is the number of discrete grid points. FFT does not require the calculation of the stiffness matrix since it is an iterative procedure [38]. The simulations converge faster than FEM for the same

geometry and resolution, while the predicted stress and strain distributions is comparable to FEM[39]. For the elasto-viscoplastic case, it was shown that the FFT solver converges much faster than FEM [40].

2.3.2.1 Crystal Plasticity simulations using FFT

Moulinec and Suquet originally proposed the use of Fast Fourier Transforms (FFT) in crystal plasticity models for simulating elastic loading in linear and non-linear composites [38]. This approach was then extended for polycrystals deforming in elastic [41], rigid viscoplastic [42] and elasto- viscoplastic [22] regimes.

To perform this simulation, the input microstructure image is discretized on a $N_1 \times N_2 \times N_3$ grid of Fourier points as seen in figure 2.5. These Fourier points correspond to the pixels in a 2D image and voxels in a 3D image. A polycrystal is represented by a periodic unit cell, as required by the FFT algorithm. The resolution of the FFT grid with respect to the synthetic microstructure is chosen according to [43] such that each grain has atleast 20 FFT points. In this work, the synthetic microstructures consisting of about 5000 grains each are discretized on a 128^3 grid, and the smallest grains have 60 or more voxels. The sensitivity analysis for FFT simulations in the elastic regime show that the stress-strain distributions are not very sensitive to the domain shape [44]. In the plastic regime, for a randomly textured material under uniaxial tension, it was shown that a grid size of 121^3 results in micromechanical fields that converge with higher grid size outputs [44]. Hence the RVE chosen for our microstructures is optimal to avoid grid size effects.

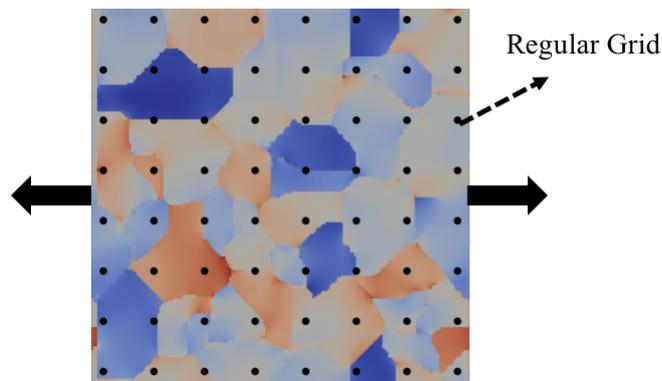


FIGURE 2.5: Representative volume element: The black dots are the Fourier points overlaid on an input microstructure image. Note that the actual grid is much finer than shown here.

At each Fourier point, the constitutive equations for stress equilibrium are solved to give the relation between Cauchy stress and strain or strain-rate tensors. The local mechanical response calculation is a simple product in Fourier space, calculated as the convolution integral between the Green's function of displacement and the heterogeneous force field. This approach solves for stress equilibrium, while satisfying the strain compatibility conditions. The local stress and strain rate responses are calculated iteratively based on an augmented Lagrangian to satisfy the governing constitutive relations at every discrete material point. FFT based methods have been developed for elastic, plastic and elasto-viscoplastic constitutive behaviors of a polycrystal. In this work we have used the elasto-viscoplastic formulation which is described in the next section.

2.3.2.2 Elasto-Viscoplastic Fast Fourier Transform (EVPFFT)

The FFT simulation takes in the microstructure image discretized on a regular Fourier grid and material properties such as the elastic stiffness moduli tensor, and slip and twinning systems for plastic loading. Note that in this work, we have ignored twinning deformation. The model gives micromechanical outputs of stress, strain-rate and orientation fields.

The FFT algorithm requires periodic boundary conditions. For experimentally measured microstructures, this requirement is generally circumvented by the use of buffer layers, which can be thought of as free space in a 3D volume, not undergoing any deformation. In this work we have used synthetically generated microstructures with periodic boundaries, thus making the use of buffer layers redundant.

The relationship between stress and strain in the elastic region at a material point x is given by:

$$\sigma_{ij}(x) = C_{ijkl}(x) : \epsilon_{kl}^e(x) \text{ in RVE} \quad (2.12)$$

where C_{ijkl} is the elastic stiffness tensor of the representative volume element (RVE). The total strain in a material is given by:

$$\epsilon(x) = \epsilon^e(x) + \epsilon^p(x)$$

where ϵ^e is the elastic strain and ϵ^p is the plastic strain in the material.

At small strains (elastic regime), the stress at any Fourier point x at time $t + \Delta t$ can be described by the constitutive equation 2.13 using Hooke's Law and an Euler implicit time discretization.

$$\sigma(x) = C(x) : (\epsilon(x) - \epsilon^{p,t}(x) - \dot{\epsilon}^p(x, \sigma) \Delta t) \quad (2.13)$$

where $C(x)$ is the elastic stiffness tensor, $\epsilon(x)$ is the total strain tensor, $\epsilon^p(x)$ is the plastic strain tensor and $\dot{\epsilon}^p$ is the plastic strain rate tensor.

At large strains (plastic regime), the elastic strains are negligible compared to the plastic strains and the following equation is solved at every material point x :

$$\dot{\epsilon}^p(x^d) = \sum_{s=1}^N m_{ij}^s(x) \dot{\gamma}^s(x) = \dot{\gamma}_0 \sum_{s=1}^N m_{ij}^s(x) \left(\frac{m_{kl}^s(x) : \sigma'_{kl}(x)}{\tau^s(x)} \right)^n \text{sgn}(m_{kl}^s(x) : \sigma'_{kl}(x)) \quad (2.14)$$

Here the summation is over all the N slip systems, τ_s , m_s and $\dot{\gamma}_s$ are the critically resolved shear stress (CRSS), symmetric Schmid tensor and local shear rate on the slip system s . $\dot{\epsilon}(x^d)$ and $\sigma'(x)$ are the strain rate and deviatoric stress tensors at the Fourier grid point x^d , $\dot{\gamma}_0$ is a normalization constant and n is the rate sensitivity exponent, which was set as 10 for all the simulations. The strain rate field is calculated such that it minimizes the average of local work satisfying the equilibrium and compatibility conditions.

For these simulations, the evolution of CRSS, τ^s with accumulated shear strain has been described using the **Voce hardening** law [45, 46] as follows:

$$\tau^s(\Gamma) = \tau_0^s + (\tau_1^s + \theta_1^s \Gamma) \left(1 - \exp\left(-\Gamma \left| \frac{\theta_0^s}{\tau_1^s} \right| \right) \right) \quad (2.15)$$

The parameters τ_0 and θ_0 refer to the initial yield stress and the initial hardening rate. $(\tau_0 + \tau_1)$ is the back-extrapolated stress and θ_1 is the asymptotic hardening rate. Γ is the accumulated shear in the grains. However, τ_s does not depend on the temperature or the loading rate. The Voce model is intended to fit a single stress-strain data set.

With an applied average strain \mathbf{E} at a strain rate $\dot{\epsilon}$, assuming periodic boundary conditions across the RVE, a small degree of homogenization is introduced in the model by creating a reference medium with an initial homogeneous reference stiffness tensor

C^o as an average of all $C_{ijkl}(x)$ values in the polycrystal as:

$$C_{ijkl}^o(x) = \frac{1}{N_{points}} \sum C_{ijkl}(x) \quad (2.16)$$

The stress is then broken up into 2 components, an average stress field, determined using the reference linear stiffness tensor; and a polarization field as :

$$\sigma_{ij}(x) = C_{ijkl}^o \epsilon_{kl}(x) + \phi_{ij}(x) \quad (2.17)$$

The polarization field ($\phi_{ij}(x)$) is the local deviation from the average stress field (σ_{ij}) at a material point x , and can be calculated iteratively using the fluctuation in stiffness and local strain [22].

The iterative elastic FFT formulation can be summarized as follows: At each time step, the local fluctuations in stress and strain are calculated iteratively [42]. First, the elastic strain at each grid point is initialized by the macroscopic average strain rate calculated from the boundary conditions at that step. Then the following steps are repeated until convergence at that time step:

- The local average stress is calculated using Hooke's law from the local elastic strain.
- The polarization field is calculated using Green's method and Fourier transform and the stress equilibrium and convergence conditions are tested.
- The stress and strain fields are updated.

2.3.2.3 Viscoplastic Fast Fourier Transform (VPFFT)

The VPFFT formulation is valid for large strains only. It assumes that the elastic strains are negligible compared to plastic strain and only the deformation in the rigid viscoplastic regime is modeled. The constitutive law in this regime between deviatoric stress and the plastic strain rate is the same as equation 2.14. Similar to EVPFFT, the plastic strain and stress fields are obtained by solving the constitutive equation iteratively at each time step, by applying the FFT algorithm. The advantage of VPFFT is that, in large plastic strain regimes, a higher strain step increment can be used in the boundary conditions, in comparison to EVPFFT. Hence to calculate final stress, strain

and orientation fields at high plastic strain, with the same accuracy, VPFFT is faster than EVPFFT.

2.3.3 Summary

Mean field models of crystal plasticity, due to the homogenization approach, cannot account for neighboring grain interactions and the intragranular heterogeneity in the micromechanical fields. Hence full field models are chosen for deformation simulations. The image based elasto-viscoplastic fast Fourier transform (EVPFFT) crystal plasticity formulation is chosen to simulate uniaxial tensile deformation in three-dimensional polycrystalline microstructures, for generating statistically significant datasets for machine learning.

Chapter 3

Machine Learning Fundamentals

Machine learning techniques are well established and have been applied to different fields such as financial services[47], [48], health care [49], [50], marketing and sales [51], [52], robotics [53], [54] and transportation [55], [56]. Machine learning (ML) is a statistical framework that automates analytical model fitting for data analysis such as finding structure in data (clustering) and making data-driven predictions or decisions. These techniques can be used to extract insights and correlations between different kinds of attributes in data.

The basic idea of machine learning is to use features in an algorithm to be fit to a target or a class. These features are individual measurable properties of the phenomenon being observed [57]. Features can be of different kinds: numeric, categorical (strings), timestamps, graph structures etc; and are processed to generate a D dimensional numerical feature vector \mathbf{X} . The feature set is then evaluated to get a reduced set of features, to improve model generalization over unseen data. Given a training dataset consisting of D dimensional N observations x_i^d , where $i = 1, 2, \dots, N$, $d = 1, 2, \dots, D$; together with the N corresponding target values y_i , the goal of machine learning is to predict the value of y for a new value of x .

Predicting stress hotspots is essentially a binary classification problem. Given a location in a microstructure, we want to predict if a stress hotspot forms there ($y_i = 1$), or not ($y_i = 0$). A feature vector \mathbf{X} whose components are derived from the microstructural descriptors is constructed. The target is a vector \mathbf{Y} having information about the

region being a hotspot. The goal of the machine learning algorithm is to fit a function $f(\mathbf{X}) = \mathbf{Y}$ which minimizes the error rate. The trained algorithm can then be used to predict the outcome \mathbf{Y} for previously unseen data.

In supervised learning i.e., when the dataset has the target \mathbf{Y} labels, it is generally split into training ($\sim 60\%$), validation ($\sim 20\%$) and test ($\sim 20\%$) sets. The models are trained on the training set and their performance is compared using the validation set. The test set is used to get an estimate of the model performance on unseen data. However, when the size of the dataset is not large, k-fold cross validation (CV) is used in practice [58]. In this technique, the training sample is randomly partitioned into k subsamples. Then $(k-1)$ subsamples are used to train the model, which is validated on the k^{th} subsample. This process is repeated k times (the folds), such that each fold is used exactly once for cross validation. The k results are then averaged to get the validation estimation. In unbalanced datasets where one of the target class population is small, stratified k-fold cross validation is used [59], where the k-folds are selected such that they have approximately equal proportion of the class labels.

There are a wide variety of ML algorithms to choose from, but can be broadly categorized into two kinds, feature based algorithms and deep learning algorithms. In this dissertation, we have used two kinds of supervised machine learning algorithms: random forest models and convolutional neural networks, which are explained in the upcoming sections.

3.1 Machine Learning algorithms

3.1.1 Decision Tree Models

Decision Tree models are suited for problems where the target variable is discrete, such as classification problems. For example, figure 3.1a shows a two-dimensional space partitioned into five regions on the basis of the two features x_1 and x_2 . A decision tree corresponding to this data partitioning is shown in figure 3.1b. At each tree node, the split is decided based on the variable that maximizes the information gain [60], [61]. Tree models are easy to interpret; by following the leaf node up the tree, all the feature contributions can be understood. However, in decision tree models, the splits in

the dataset are aligned with the axes, and hence large tree depths are required to fit a generic decision boundary. Also, each leaf node in the tree is associated with a single target value, and hence these models are prone to overfitting the training dataset. Hence we switch to "tree ensemble learning" models that generate many classifiers and aggregate their result. We have used a random forest model in Chapter 4 and 5, which is explained in the following section.

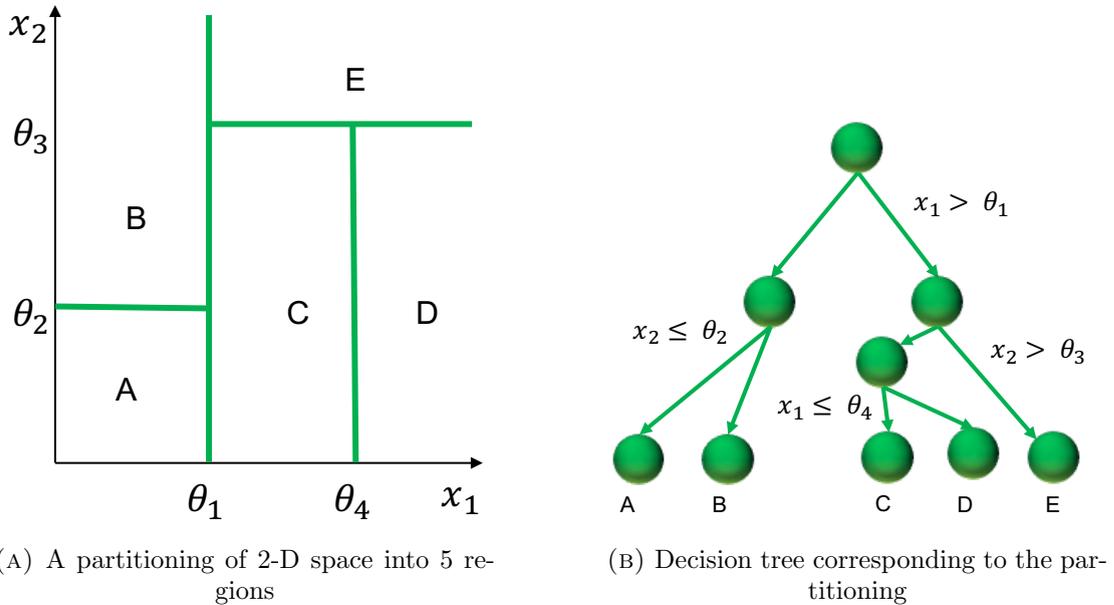


FIGURE 3.1: Schematic of a decision tree model

3.1.2 Random Forest Tree Models

The random forest (RF) algorithm is built on the concept of decision trees [62], [63]. These models are very fast and easy to fit, can handle both numerical and categorical features and deal with missing or unbalanced data efficiently. The idea behind ensemble learning is that combining weak learners results in a "strong learner". A decision tree in itself is a weak model. Random forest utilizes ensembling to bring together a number of weak decision trees. The schematic for random forest algorithm is shown in figure 3.2. The training data is bootstrapped (sampling with replacement) into n random subsets. A decision tree is fit to each bootstrapped sample using a small random subset of the feature space. This way 200-500 decision trees are grown, hence the name random forest. The output from each of the decision trees is then voted or averaged to get a final prediction. RF models prevent overfitting by using random subsets for both feature

selection and data division. The RF algorithm was implemented in Python using the

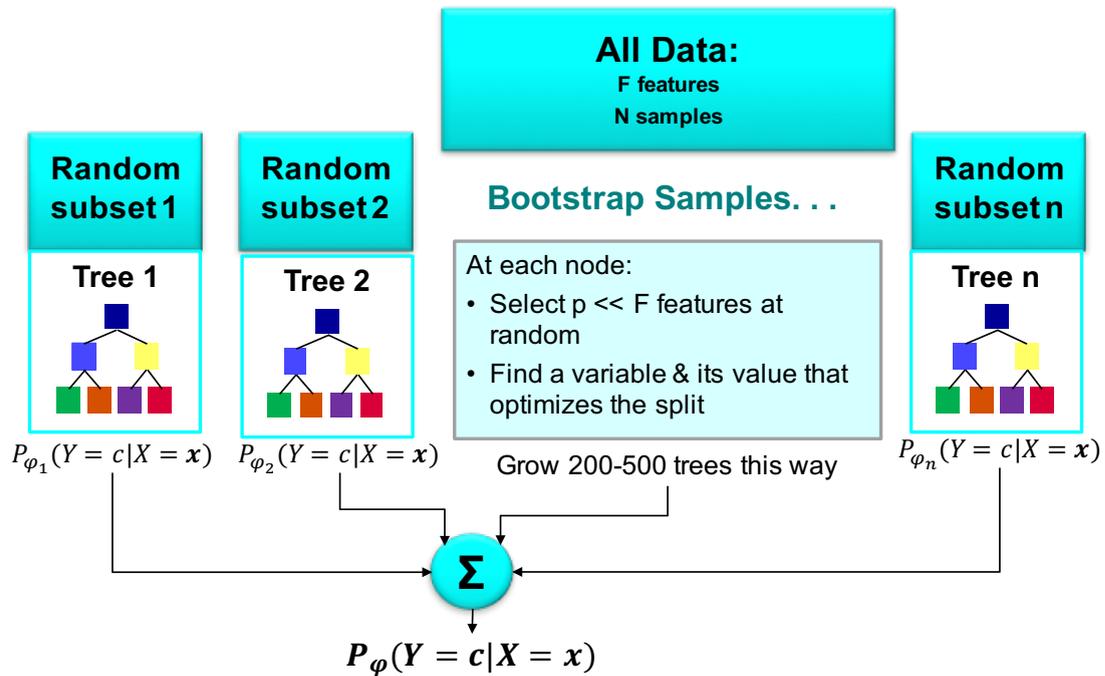


FIGURE 3.2: Schematic of the random forest algorithm.

scikit-learn implementation [64]. Some of the important model hyper-parameters for RF are:

- `max_features`: The number of features to consider when looking for the best split for a decision tree. This parameter prevents overfitting. A rule of thumb is to keep it smaller than \sqrt{D} , where D is the dimension of the feature vector.
- `num_estimators`: The number of decision trees in the forest. This parameter needs to be monitored on a neutral test set to prevent overfitting.
- `max_depth`: The maximum depth of each decision tree. Increasing the tree depth beyond an optimum will result in overfitting.
- `random_state`: This parameter can be used to seed the random number generator, so that the same random subsets are generated.

The model is trained on a set of hyperparameters, and the model performance on a test dataset is compared to choose the best performing model hyperparameters. The RF model outputs the probabilities of the sample to belong to a class, which can then be thresholded to get predicted labels.

3.1.2.1 Variable Importance

The RF algorithm can be used to determine variable importance, i.e. which of the input features most influences the target. The algorithm is such that it captures the predictors which become important due to their interactions with other variables. The random forest algorithm estimates the importance of a variable by comparing the increase in the prediction error when the out-of-bag [65] data for that variable is permuted, while all others are left unchanged [63]. Out-of-bag is the mean prediction error on each training sample x_i , using only the trees that did not have x_i in their bootstrap sample.

For each of the trees in the forest, a predictor variable is permuted while all others are left unchanged. Variable importance is estimated by looking at the difference in prediction accuracy before and after permuting that feature. By randomly permuting the predictor variable X_j , its original association with the response Y is broken. When the permuted variable X_j , together with the remaining un-permuted predictor variables, is used to predict the response, the prediction accuracy (i.e. the number of observations classified correctly) decreases substantially, if the original variable X_j was associated with the response. Thus, a reasonable measure for variable importance is the difference in prediction accuracy before and after permuting X_j . The variable importance calculated this way is known as “Permutation Accuracy Importance (PAI)”. PAI along with correlation coefficients can be used to gain insights into the physical problem. Another application of PAI is the feature selection process, where we can discard the predictor variables with the lowest PAI values. This model reduction helps in making the models simpler and more interpretable.

3.1.3 Deep learning

Deep learning methods have recently outperformed many state of the art machine learning techniques in several fields such as computer vision[66], speech recognition[67], medical image segmentation[68], [69], [70], and video analysis [71] and require minimal feature engineering by hand. Deep learning [55] utilizes neural networks to fit a function to the dataset. It refers to biologically-inspired computational models that are “composed of multiple processing layers to learn representations of data with multiple levels of abstraction” [55]. These models have outperformed previous state-of-the-art tasks in different

domains, which has led to a surge of interest in their applications. These methods have found their way into the materials science domain, for automatic microstructure classification [72], [73], predicting material properties from crystal structures [74], [75] and predicting flow stress behaviour [76], [77]. Neural networks approximate a function with a densely interconnected sets of simple function units (e.g. sigmoid). Each unit takes real valued inputs, this input could consist of the output of other units. The unit then produces a single real-valued output which might become input of many other units. These models are incredibly complex and can produce highly non-linear decision boundaries.

3.1.3.1 The Perceptron

The perceptron is the basic building block of neural networks. Figure 3.3b shows a single layer perceptron, which is the simplest neural network possible. The output is obtained from a series of functional transformations. First, a linear combination of the inputs (\mathbf{x}) is created as:

$$a(\mathbf{x}, \mathbf{w}, \mathbf{b}) = \sum_{i=1}^4 (w_i x_i + b_i)$$

where \mathbf{x} is the input data, \mathbf{w} is the weights matrix, and \mathbf{b} is a bias or threshold value that helps define the activation function, $a(x)$, which is a linear classifier of the input data. In order to learn non-linear decision boundaries, it is necessary to transform the activations $a(x)$ through a non-linear function activation function $f(x)$ such that the output of the neuron $h(x)$ is obtained as:

$$h(x) = f(a(\mathbf{x}, \mathbf{w}, \mathbf{b}))$$

Common choices of the non linear activation function are sigmoid function, rectified linear unit (ReLU), leaky ReLU and hyperbolic tangent (tanh) functions, are shown in Figure 3.3b. The output of the single layer perceptron is:

$$y(\mathbf{x}, \mathbf{w}, \mathbf{b}) = \begin{cases} 1, & \text{if } h(x) > 0. \\ 0, & \text{otherwise.} \end{cases} \quad (3.1)$$

This process is known as *forward propagation* of information through the network. The weights \mathbf{w} and biases \mathbf{b} are optimized through supervised learning via *backpropagation*

using a gradient descent algorithm to minimize a loss function. The loss function represents the "cost" associated with the difference between true labels and predicted output, which can be a simple mean squared error function for regression tasks, or a cross entropy error for classification tasks. The network training happens over a series of iterations known as *steps*, where the weights are updated based on the backpropagated gradients. An *epoch* of training is completed after the network backpropagates gradients for the entire training data for optimizing the weights. Figure 3.4 shows how stacking multiple perceptrons in a layer and stacking layers on layers, perceptrons can be constructed into deep architectures known as multi-layer perceptrons or artificial neural networks. The four nodes in the left side of the cartoon represent the input data. The middle layers consisting of 7, 3 and 2 units respectively represent the three hidden layers of the network. The output layer of this network has one unit. The number of units and number of layers can be varied to build a complex neural network. The reader is referred to Lecun et al. [55] for a detailed introduction to deep learning methods.

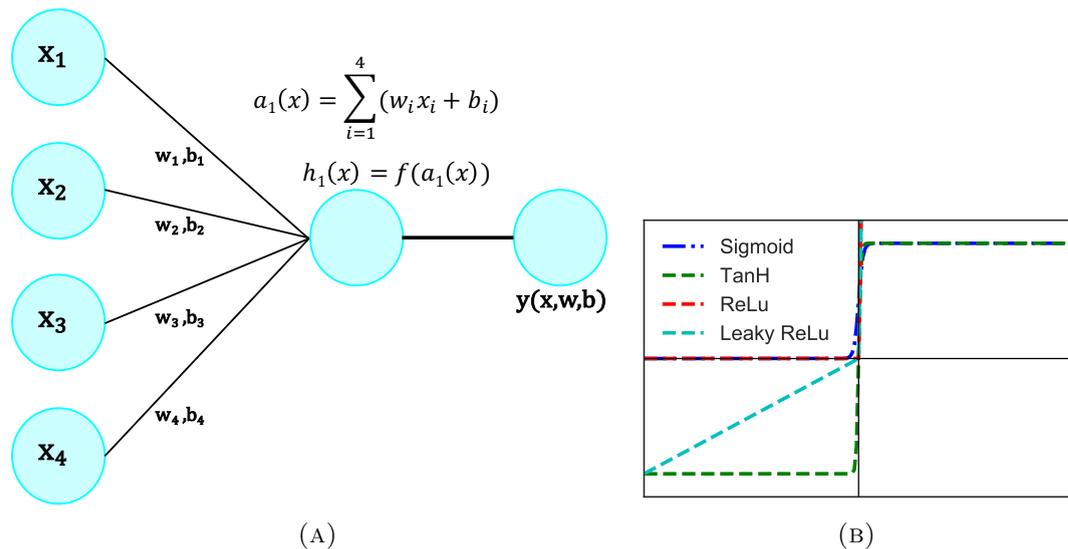


FIGURE 3.3: a) Schematic of a single layer perceptron b) Different activation functions in a perceptron

3.2 Model performance metrics

The test for a machine learning model is the validation error on new data. In this work, predicting stress hotspots has been treated as a classification problem. The predicted probabilities are thresholded and the labels are assigned. To choose the best threshold,

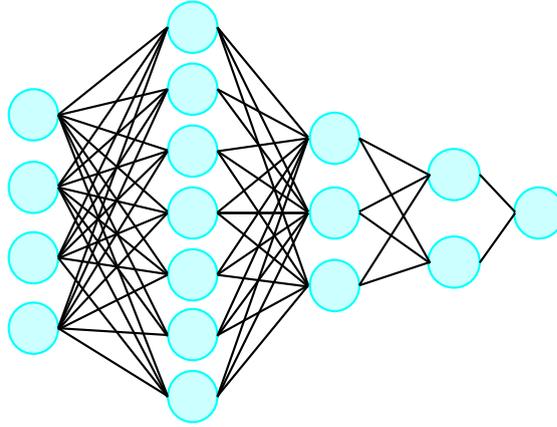


FIGURE 3.4: Schematic of an artificial neural network constructed from perceptron units

model performance needs to be compared. Some of the common performance measures for such models are explained in this section.

3.2.1 Accuracy, Precision and Recall

Accuracy is the proportion of correct results achieved by a classifier. However, in imbalanced datasets, for example if only 10% of the dataset has the label 1, then predicting all the labels as 0's will still result in a 90% accuracy. This is because accuracy assumes equal cost for false positive and false negative errors. Thus precision and recall are used to compare model performance.

$$Precision = \frac{TP}{TP + FP} \quad (3.2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3.3)$$

where TP, FP, TN, FN are the number of true positive, false positive, true negative and false negatives for a set of predictions. Precision tells us the fraction of correctly classified points out of those predicted to belong to a class. Recall tells us the fraction of correctly classified points out of the total number of points in that class. These measures can be used to compare model performance for each threshold value. In order to compare the model performance across the range of probability thresholds, the area under curve of receiving operator characteristic curve (AUC) metric is used, which is explained below.

3.2.2 Area Under Receiving Operator Characteristic Curve (AUC)

The receiving operator characteristic (ROC) curve is a plot of true positive vs false positive rate, across the probability thresholds. Each probability threshold represents different trade-off between FP and FN. The cost ratio is the slope of the ROC curve, and the area under the curve represents the performance averaged over all thresholds for a classifier. If the classifier is very good, the true positive rate will increase quickly and the area under the curve will be close to 1. If the classifier is no better than random guessing, the true positive rate will increase linearly with the false positive rate and the area under the curve will be around 0.5.

AUC is independent of the fraction of the test population which is class 0 or class 1 and is hence useful for evaluating the performance of classifiers on unbalanced data sets. If the ROC curves of 2 classifiers intersect, it means that one of them is better for some cost ratios. If the ROC curves of 2 classifiers do not intersect, the classifier with higher AUC is better. Figure 3.5a shows the schematic of an ROC curve as the threshold value for classifying a point label as 0 or 1 is increased from 0 to 1. In this work, the trained models are evaluated using 5-fold cross validation AUC score.

Another metric for measuring model goodness in imbalanced datasets for binary classification is the Matthews correlation coefficient (MCC) as it takes all the elements of the confusion matrix (true and false positives and negatives) into account [78]. However, this measure has not been used in this work.

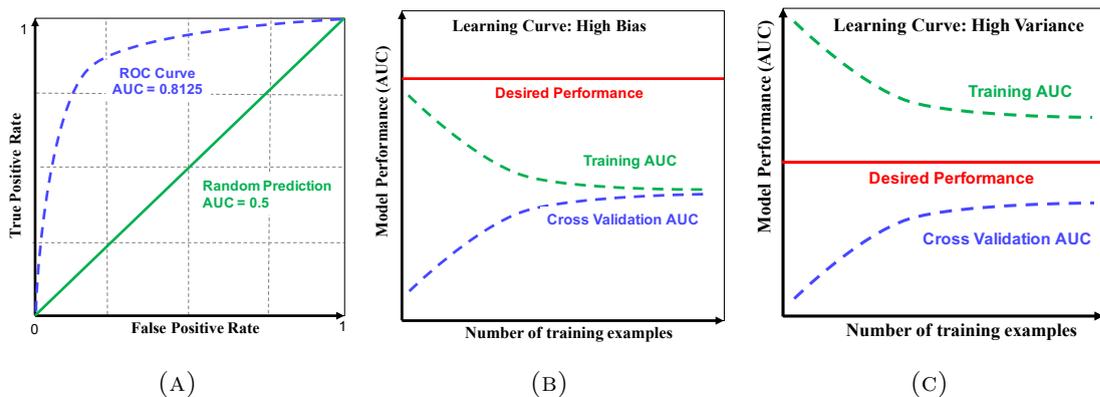


FIGURE 3.5: Schematic of 3.5a: the ROC curve and AUC metric. The diagonal line represents random prediction, 3.5b: learning curve for the high bias case and 3.5c: learning curve for the high variance case

3.3 Model Generalization

The prediction error of a model can be decomposed into three kinds of errors: bias (underfitting), variance (overfitting) and irreducible errors arising from noise in the data. During supervised learning, understanding the trade-off between bias and variance helps in diagnosing the models and generalizing them beyond the training dataset. The ability of the model to approximate the data is related to bias, hence, if the model is not complex enough, it will result in a bias or under-fitting the data. For example, trying to model a quadratic relationship using a linear model results in an a prediction error related to model bias. Variance is related to the sensitivity of the model to the training dataset. A model with high variance overfits the training data to learn the random noise, and hence performs worse on unseen test data.

Learning curves are diagnostic curves that can be used to improve model performance by understanding the kind of error it is suffering from. Learning curves are a plot of the prediction score on the training and cross validation datasets as a function of the training dataset size. In models with high bias error, the cross-validation (CV) performance initially increases, but then plateaus as the number of training examples increases; the training performance initially decreases, but then plateaus at the level of CV performance, which is usually lower than the desired performance level. Thus in learning curves of high bias models, the training and CV performances are similar for a range of training set sizes but lower than the desired performance. Expanding the feature space helps in correcting models with high bias error. In models with high variance error, the training performance decreases with number of training examples, usually to a level higher than high-bias case. The CV performance increases with the number of training examples, but there is a gap between training and CV performances. Increasing the size of the training dataset or using a smaller feature set helps in correcting models with high variance error. Figure 3.5c and 3.5b show a schematic of how learning curves can be used to diagnose model errors.

Chapter 4

Predicting Stress Hotspots in Face Centered Cubic Materials

We investigate the formation of stress hotspots in polycrystalline materials under uniaxial tensile deformation by integrating full field crystal plasticity based deformation models and machine learning techniques to gain data driven insights about microstructural properties. Synthetic 3D microstructures are created representing single phase equiaxed microstructures for generic copper alloys. Uniaxial tensile deformation is simulated using a 3-D full-field, image-based Fast Fourier Transform (FFT) technique with rate-sensitive crystal plasticity, to get local micro-mechanical fields (stress and strain rates). Stress hotspots are defined as the grains having stress values above the 90th percentile of the stress distribution. Hotspot neighborhoods are then characterized using metrics that reflect local crystallography, geometry, and connectivity. This data is used to create input feature vectors to train a random forest learning algorithm, which predicts the grains that will become stress hotspots. We are able to achieve an area under the receiving operating characteristic curve (ROC-AUC) of 0.74 for face centered cubic materials modeled on generic copper alloys. The results show the power and the limitations of the machine learning approach applied to the polycrystalline grain networks.

4.1 Introduction

Ductile fracture is one of the most common modes of failure in materials and occurs by the nucleation, growth and coalescence of microscopic voids. [17] established that these

voids grow under stress by accumulating vacancies, and that void nucleation is induced by stress. [20] show that an applied stress on a material is heterogeneously distributed between the grains, creating regions of stress accumulations, so-called stress hotspots. Stress distribution between grains is dependent on the local microstructural features which in turn influence the location of void nucleation. We propose using machine learning techniques to study the impact of microstructural features on stress hotspots. Predicting damage nucleation is important because fracture ultimately defines the useful lifetime of a material.

Modern texture analysis techniques, such as near field and far field High Energy X-Ray Diffraction Microscopy (nf-HEDM, ff-HEDM) [79], have made three dimensional characterization of microstructures possible. This kind of mesoscale microstructural data, consisting of grain crystallography, centroids and strain fields, is well suited to machine learning techniques. Following this trend, [80] have recently used decision trees to find insights about the driving forces behind deformation twinning in a magnesium alloy. However, stress hotspots are rare events; consisting of less than 10% of the material volume and hence require a large dataset for statistical learning, which is not amenable to the currently existing HEDM datasets. Instead, we meet this requirement by using a simulation generated data set. Uniaxial tensile deformation is simulated in a number of synthetic microstructures using an image based full field crystal plasticity Elasto-Viscoplastic Fast Fourier Transform (EVPFFT) model [22]. Simulating deformation in materials gives us the advantage of preserving both the initial and final structures, thus enabling us to turn back the clock and correlate hot spots to initial microstructural features. The techniques developed in this work are directly transferable to HEDM datasets.

With the advent of material informatics, machine learning has been used to search for stable compounds across composition space, extract correlations between physical characteristics and observed properties and search for materials with useful properties [80], [9]. Taking inspiration from these successes, we use data mining and machine learning techniques at the mesoscale to build models which can be used to predict the probable failure locations based on known data in similar materials. These models help us gain insights about the characteristics in the local structure such as texture and geometry, that allow stress hotspots to form, and how such regions can be identified in a microstructure.

In this chapter, plastic deformation of single phase face centered cubic (FCC) polycrystals is studied to ascertain the local microstructural characteristics related to the regions of high stress concentrations. A random forest learning algorithm is chosen for its ease of use and interpretability of the machine learning framework. We use statistical descriptors of microstructures describing the crystallography (orientation distribution function, Schmid factor, misorientations) and geometry (grain shape, grain boundary types) which have not been used earlier to understand their correlation with hotspot locations. The objective is twofold: predict probable failure locations, and identify the microstructural features that cause them, to facilitate microstructure engineering for materials design.

4.2 Methods

4.2.1 Dataset Generation

4.2.1.1 Synthetic Microstructures

First, a dataset of synthetic microstructure images is built using Dream.3D [81]. Synthetic equiaxed polycrystalline microstructures with about 5000 grains each and a mean grain size of 2.7 microns are created for a set of representative textures in FCC materials shown in Figure 4.4a. These textures were selected from a number of common FCC rolling texture components such as brass, copper and Goss components. For each texture kind, six microstructure instantiations are created, thus resulting in about 30000 grains per crystallographic texture. The texture intensity is characterized by multiples of random density (MRD) which is the intensity of a crystallite orientation with respect to its intensity in a randomly textured material. For each texture studied here, the texture intensity is varied from weakly textured (<10 MRD) to strongly textured (>30 MRD). The microstructures are discretized on a $128 \times 128 \times 128$ grid, which allows the use of image based crystal plasticity models.

The stress distribution in a microstructure is highly dependent on the crystal system and the texture. Hence we keep the crystal system constant (FCC) and vary the texture, while keeping the grain size distribution, slip system strength, slip hardening rates, strain rate and other factors constant while simulating uniaxial tensile deformation.

4.2.1.2 Simulating Uniaxial Tensile Deformation

Due to the homogenization approach, mean field models of crystal plasticity cannot account for neighboring grain interactions and the intragranular heterogeneity in the micromechanical fields. Hence a full field elasto-viscoplastic fast Fourier transform (EVPFFT) crystal plasticity formulation [22] is chosen to simulate uniaxial tensile deformation in the generated microstructures. This model takes in the microstructure image discretized on a regular Fourier grid and the material properties such as the elastic stiffness tensor and slip systems for plastic loading. The output is the stress, strain-rate and orientation fields at each grid point. Since the EVPFFT model solves the constitutive equations at each grid point, the grain size should not affect simulations as long as the Fourier grid size is chosen such that the model converges. The boundary conditions on the strain are chosen such that the material transitions into the plastic regime.

The Voce hardening law [45] is used to model the evolution of the critically resolved shear stress (CRSS) of each slip system s , τ^s , with accumulated shear strain as follows:

$$\tau^s(\Gamma) = \tau_0^s + (\tau_1^s + \theta_1^s \Gamma) \left(1 - \exp\left(-\Gamma \left| \frac{\theta_0^s}{\tau_1^s} \right| \right) \right) \quad (4.1)$$

The parameters τ_0 and θ_0 refer to the initial yield stress and the initial hardening rate. $(\tau_0 + \tau_1)$ is the back-extrapolated stress and θ_1 is the asymptotic hardening rate. Γ is the accumulated shear in the grains. The Voce hardening parameters were extracted as shown in table B.1 by fitting the VPSC simulated stress-strain curve to representative experimental stress-strain curves for FCC copper. Appendix A covers the details of extracting these parameters.

4.2.1.3 Problem Formulation: defining stress hotspots

EVPFFT simulations result in a voxel-wise output for the Von Mises (VM) stress field, as shown in Figure 4.1a. It is observed that the regions of high stress generally form in clusters and intra-grain variations in stress values are small. Therefore, to minimize the impact of numerical artifacts and small-scale fluctuations, this field is averaged grain-wise to get the stress in each grain. The resultant stress distribution is thresholded using the peak over threshold method [82] to select the critical stress threshold (Figure 4.1b). The grains having VM stress above the critical threshold are designated as stress hotspots. The critical threshold grain averaged Von Mises stress value was found to lie

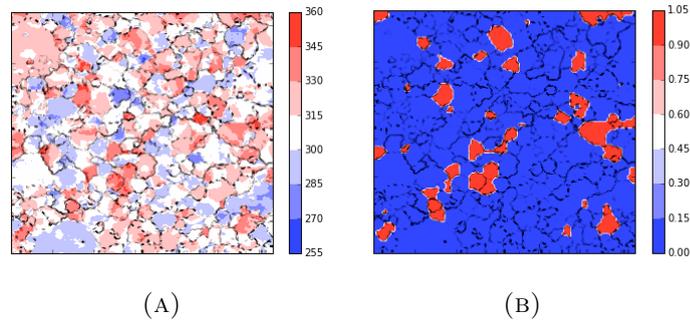


FIGURE 4.1: Cross sections of a microstructure showing (a) Von Mises Stress field and (b) Stress hotspots obtained by thresholding on grain averaged stress field. The red regions correspond to regions of high stress or stress hotspots.

between the 85th and 95th percentile. Hence the 90th stress percentile was chosen as a cutoff throughout the dataset to keep the fraction of hotspots the same between all the microstructures.

4.2.1.4 Relation Between Stress Hotspot locations

A spatial autocorrelation analysis was performed to determine if the location of a hot grain has a role in the formation of another hotspot, i.e. if the hot grains are related to each other. Two-point statistics are calculated between hot grains and normal grains using the method described in [83]. We observe a high autocorrelation inside the coherence length which represents the average size of the hotspots. Outside this length, we find that the probability of finding a hotspot is constant in all directions, which means that the hot grains are not correlated in space i.e. they are dispersed uniformly in the microstructure. Thus stress hotspots (grains) are independent of each other and standard statistical approaches can be used without having to worry about the shadowing effects of spatial correlation.

4.2.1.5 Effect of Microstructure evolution on Stress hotspots

As a material deforms, the grains tend to rotate towards the tensile axis, thus changing their orientation. [22] have used the EVPFFT formulation to study hotspots formed during uniaxial tensile deformation in polycrystalline FCC materials. It was observed that in materials where the "hard" and "soft" directions in the elastic and plastic regimes are different, the elastic hotspots become plastic cold-spots and vice-versa.

The constitutive model parameters for FCC materials represent a generic copper alloy. For the single crystal elastic constants used in this work, the elastic anisotropy

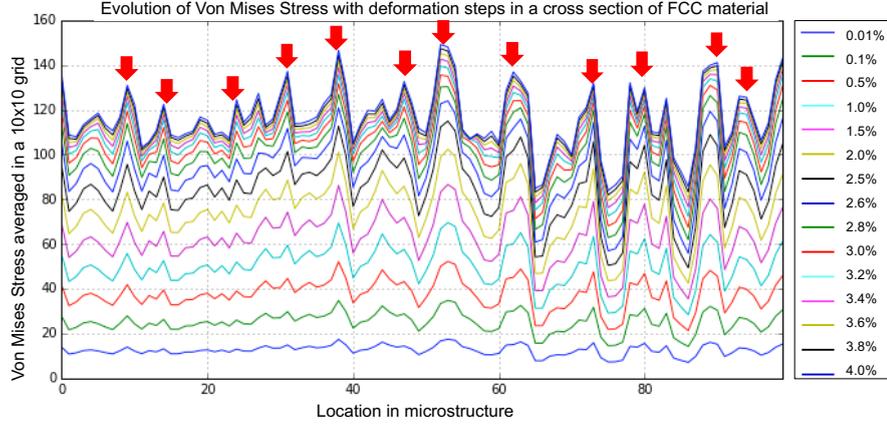


FIGURE 4.2: Representative figure showing the Von Mises stress vs. location in the microstructure at each deformation step. The bottom (blue) curve is at the lowest strain (0.01%) and the top (blue) curve is at the maximum strain (4%). The red arrows indicate how location of stress hotspots is not changing

parameter given by equation 4.2 ([84]) is $A = 3.2$.

$$A = (2 \times C_{44}) / (C_{11} - C_{12}) \quad (4.2)$$

where C_{44} , C_{11} and C_{12} are elements of the elastic constant tensor of a cubic material. For $A > 1$; $\langle 100 \rangle$ and $\langle 111 \rangle$ are soft and hard elastic directions respectively. This coincides with the plastic anisotropy of $\{111\} \langle 110 \rangle$ slip ([22]), not taking the effect of strain rate into account. Hence, elastic hotspots should remain plastic hotspots (and cold spots remain cold) as the deformation proceeds.

To verify this, the location of stress hotspots in the elastic and plastic regimes was studied. A 128×128 cross section of the microstructure was divided into 196 10×10 regions. The Von Mises stress field was averaged in each 10×10 region, and is plotted at each deformation step in Figure 4.2. It was found that the stress hotspots are stationary in the crystal structures studied. That is, initial high stress regions remain high in stress as deformation proceeds; likewise, low stress regions retained low stress values throughout deformation.

4.2.1.6 Effect of Texture and Microstructure

We conducted two experiments. First we compared the location of stress hotspots in microstructures having the same grain structure but different texture templates applied to it. In the second experiment, we compared the location of stress hotspots in microstructures having similar random textures, but stochastically different microstructures. All

the microstructures have equiaxed grains and a lognormal grain size distribution with mean grain size of 2.7 microns. We found that changing either the texture or the grain structure has an impact on the macroscopic flow stress-strain response and the location of hotspots.

4.2.2 Developing Microstructural Descriptors

Feature engineering is the process of using knowledge of the pertaining field (in this case, deformation mechanics) to build features to be used by a machine learning algorithm [85]. During material deformation, the loading condition and microstructure are the most important factors that determine stress distribution in a microstructure. For stress hotspot prediction, both crystallographic and geometric microstructural descriptors form the domain knowledge based feature set. Crystallography and geometry based microstructural representations to be used as features during machine learning are developed in the next sections. Table A.3 lists the acronyms and descriptions of the features used in this work.

4.2.2.1 Crystallographic Descriptors

The shear strain rate on a material point depends on the tensor dot product between the Schmid tensor and the deviatoric stress tensor. These tensor quantities are determined by the grain orientation, which is captured by the Euler angles [24], [86]. However, due to the non-linearity of Euler space, there can be multiple Euler angles representing the same orientation. Interpreting the results from Euler angle ranges associated with hotspots is difficult even after reducing the Euler angle space to the fundamental zone. The complex, non-linear, trigonometric relationship between the tensor quantities cannot be captured directly by using Euler angles as features, and hence we develop the following features to represent the grain crystallographic properties.

Distance from Inverse pole figure corners: A microstructural descriptor based on the inverse pole figure (IPF) space is proposed to capture the relation between the tensile axis direction and the grain soft and hard orientations resulting from the anisotropy in elastic modulus. The loading direction is projected in the inverse pole figure space in the fundamental zone. This ensures that every orientation is projected into the same stereographic triangle. The distance of each projected point from the three corners of the inverse pole figure describes the orientation of the loading direction

w.r.t. the 3 crystal directions. Since this space is fixed, the Euclidean distance of the projected point from the 3 corners of the inverse pole figure is invariant and is used as a microstructural descriptor. This distance is an indicator of how close the tensile axis is to the crystal directions in the inverse pole figure. Figure 4.3a shows a schematic of the three distances in cubic materials. The features derived would be 001_IPF_0 : distance of the $[001]_{sample}$ direction from $[111]_{crystal}$ direction; 001_IPF_1 : distance of $[001]_{sample}$ from $[001]_{crystal}$ and finally, 001_IPF_2 : the distance of $[001]_{sample}$ from $[101]_{crystal}$ direction. Similarly, other directions in the sample frame were projected into the crystal frame to get corresponding features of the form $SampleDirection_IPF_x$, where $x=\{0,1,2\}$.

Misorientation: [87] show that during aluminum bicrystal deformation, the yield stress and the rate of work hardening increases with the orientation difference between the crystals. Some misorientations correspond to "special" grain boundaries which could lead to enhanced properties with respect to corrosion, impurity segregation, cracking, coarsening, diffusion and other properties affected by grain boundary properties [88]. Hence, the misorientations of grains with their 1st, 2nd and 3rd set of nearest neighbors were calculated, and features like minimum, maximum and mean misorientation for a grain were calculated from these lists. Figure 4.3b shows a schematic 2-D microstructure with a grain, its 1st nearest neighbors and the average misorientation calculation. The misorientation with the 1st nearest neighbors in 3D microstructures was calculated using the "Find Feature Neighbor Misorientations" filter in Dream.3D.

Slip Transmission Metrics: A geometric compatibility factor to measure the ease of slip transmission between two grains is the mprime factor from [89], calculated as a dot product between the slip plane normals and slip directions across a grain boundary as seen in figure 4.3c. The m-prime factor lies in the range (0,1); a value of 0 indicates incompatible deformation across the grain boundary and a value of 1 indicates co-planarity of slip systems between two grains. [90] show that this factor is related to the local misorientation between the slip systems in different grains and can help in predicting the slip transmission across different grains. This metric is calculated using a filter in Dream.3D for FCC materials. The mprime factor is calculated between a grain and all its contiguous neighbors. This list is averaged to get a single mprime value across the grain boundaries of a grain.

Schmid Factor: The Schmid factor is a measure of the optimal orientation of a slip system for deformation in a single crystal, and can be extended to polycrystalline

materials [86].

Other Factors The *Taylor factor* [91] consists of information about the slip occurring in each grain and the equivalent Von Mises stress and is calculated during deformation. Since our goal is to correlate only the initial microstructural descriptors to stress hotspots, the *Taylor factor* is not included in the feature set. Factors like *stress triaxiality*, *slip system activities*, *principal stress directions* and *Lode angle parameters*, are important in determining the stress state in a grain, but are dropped from the feature space for similar reasons.

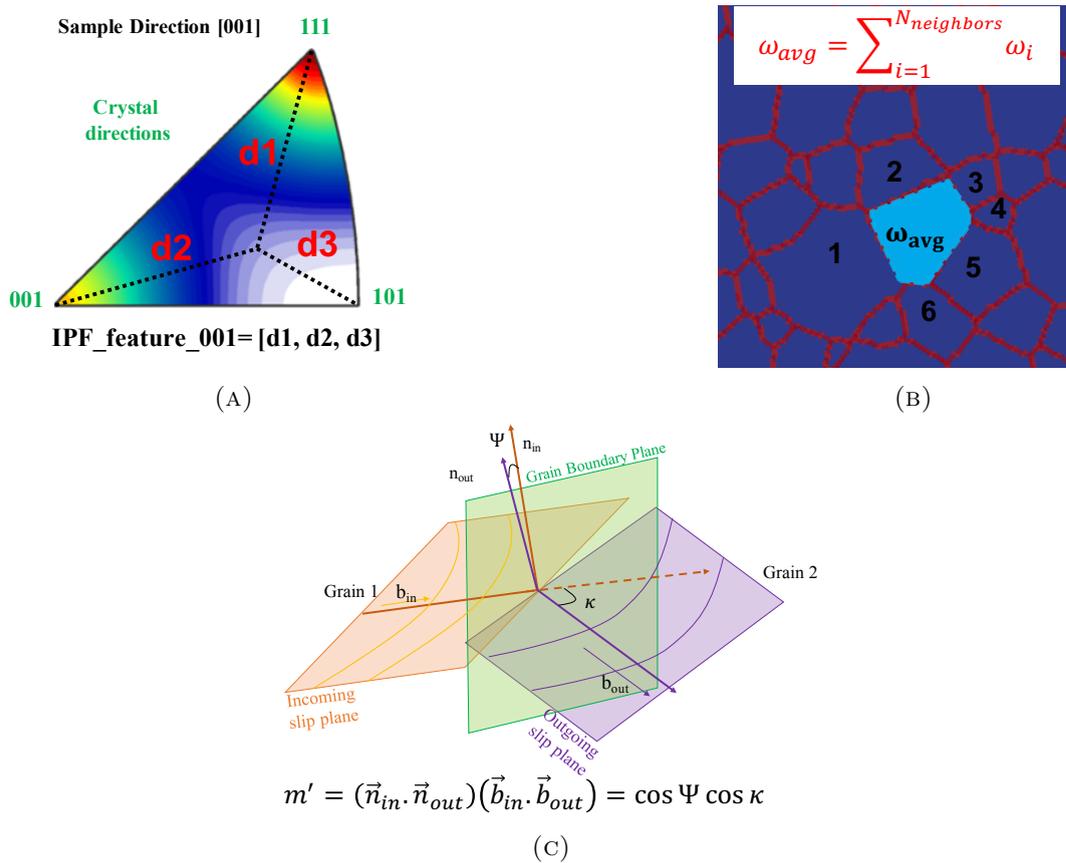


FIGURE 4.3: Crystallographic descriptors used as features during machine learning. (a) Schematic of an Inverse pole figure for a cubic crystal. The distance of the [001] sample direction from the three crystal directions [001] (d2), [101] (d3) and [111] (d1) in the standard stereographic triangle are used as a feature. (b) Schematic of average misorientation ω_{avg} for the reference (light blue) grain with respect to its N contiguous neighbors (numbered 1 to 6) in a 2-D microstructure. ω_i is the misorientation between the reference grain and its i^{th} neighbor grain. (c) Slip Transmission Factor (m' , m') is a function of the dot product of the slip plane normals (incoming normal n_{in} and outgoing normal n_{out}), and the dot product of slip directions (incoming d_{in} and outgoing d_{out}). Adapted from [1]

4.2.2.2 Geometrical Descriptors: Grains and their neighborhood

Material failure is highly dependent on microstructure. For example, cleavage fracture in mild steels has been shown to have a grain size dependence [92], whereas during ductile fracture under uniaxial tensile deformation, stress hotspots tend to form near microstructural features and usually form in textures corresponding to maxima in Taylor factor [21]. In this section, we develop a few microstructural descriptors describing the grain geometry which are used in this work.

Shape averaged distance from special points: Due to the three-dimensional polycrystal line grain structure, the grain boundary network consists of grain boundaries where two grains meet; *triple lines*, where 3 grains and 3 boundaries meet; and *quadruple points*, where 4 grains, 6 boundaries and 4 triple lines meet [93]. The grains closer to these special points have a neighborhood with a higher local heterogeneity. The triple junctions, depending on their crystallography, can impede or permit inter-granular damage through them [94], and hence are an important microstructural descriptor. Rollett et. al.[21] show that stress hotspots tend to form in grains closer to these special points. Hence we use the average distance of each grain to the closest grain boundaries, triple lines, and quadruple points as geometric features. A Dream.3D filter is used to calculate the Euclidean distances for each voxel, which are then averaged per grain to form a grain level feature.

Grain shape parameters: It has been shown that in severely plastically deformed (SPD) materials, such as those produced by high pressure torsion (HPT), the ductility is influenced by the grain aspect ratio [95]. In alpha-beta Titanium alloys, the grain size, aspect ratios and grain size distribution have been shown to be related to the flow stress [96]. The Hall-Petch and inverse Hall-Petch relationships describe the relation between yield stress and grain size in a material [97]. In this work, we have used equivalent spherical grain diameter and grain aspect ratio from a best fit ellipsoid as features. They were calculated using a Dream.3D filter. The grains can also be characterized by the number of contiguous neighboring grains, which is used as a feature and calculated via Dream.3D.

4.2.3 Machine Learning Methods

The EVPFFT simulations and the derived microstructural descriptors form a sample on which machine learning models can be built and evaluated. The microstructural descriptors are used as features to train machine learning algorithms for predicting stress hotspots. The contribution of each microstructural feature in predicting hotspot formation is studied to understand its impact. Predicting stress hotspots is a binary classification problem. Given a grain in a microstructure, we want to predict if a stress hotspot forms there. A feature vector \mathbf{X} whose components are derived from the microstructural descriptors is constructed. ML methods can be used to extract insights and correlations between the elements of \mathbf{X} and the micromechanical outcome.

4.2.3.1 Model Performance Metrics

Once a model is trained, its performance can be evaluated by looking at the classification accuracy. However, for stress hotspot classification, predicting all the grains as normal (non-hotspots) will still result in a 90% classification accuracy as only 10% of the grains are hotspots. A better representation of the classification accuracy is the two-dimensional confusion matrix, indexed in one dimension by the true class and in the other by the predicted class. A correctly predicted hot grain is true positive; a misclassified hot grain is a false negative. The normal grains predicted as hot are false positives, and the normal grains predicted normal are true negatives. We want the classifier to correctly predict all hotspots, maximizing the true positives while minimizing the false positives. Hence the area under the receiving operator characteristic curve (AUC) is a good evaluation metric for such unbalanced datasets [98]. The area under the receiving operator characteristic curve (AUC, described in section 3.2.2) is a good evaluation metric for such unbalanced datasets [98]. If the classifier is no better than random guessing, the true positive rate will increase linearly with the false positive rate and the area under the receiving operating characteristic curve will be around 0.50. A good classifier has a high true positives rate and a low false positive rate, and the $AUC \sim 1$. A misclassified hot grain (false negative) has a higher cost and hence the recall is important for hotspot classification.

4.2.3.2 Estimation of Model Generalization Error

A data-driven model should achieve good performance on the training data as well as generalize well on unseen (test) data. We divide our dataset into training and test (holdout) sets. During training, the model parameters are optimized using K-fold cross validation (CV) ([99]). In this technique, the training sample is randomly partitioned into k subsamples. Then $(k-1)$ subsamples are used to train the model, which is validated on the k^{th} subsample. This process is repeated k times (the folds), such that each fold is used exactly once for cross validation. The k results are then averaged to get the validation estimation. To offset the effect of unbalanced data, stratified k-fold cross validation is common ([59]), where the k-folds are selected such that they have approximately equal proportion of the class labels. In this work, the grain-wise hotspots are designated based on the crystal plasticity simulations and might be correlated within a microstructure as they come from the same crystal plasticity simulation. Therefore, to assess the generalization error during validation and test times, we perform stratified sampling. The validation folds are created by selecting the grains from a randomly chosen microstructure per texture class that is absent in the training data. This overcomes the optimistic bias in generalization error from having correlated data from a single simulation between training and validation. Once the model hyper-parameters are optimized, a second K-fold cross validation is run by using the entire training dataset to construct the models and creating validation folds from the microstructures in the holdout (test) dataset. This helps in getting an estimate of the generalization error on unseen data. This technique is also known as nested cross-validation ([100]).

The generalization error can be understood by dividing it into bias and variance errors. If the learning algorithm is too simple, it introduces a bias in the predictions, whereas a very complex algorithm can overfit and learn the noise in the dataset, leading to a higher variance in the predictions. Learning curves are diagnostic curves that can help in understanding the trade-off between bias error and variance error, thus helping the models to generalize beyond the training dataset. They are a plot of the model evaluation metric on the training and cross validation datasets as a function of the training dataset size.

The model performance generally increases with the training dataset size. If there is a gap between training and validation error/ performance, the model is suffering from high variance and collecting more data will help. On the other hand, the model

suffers from bias when training and testing errors converge and are high, and a more complex algorithm or more features are needed. These curves can be used to determine the size of training dataset required.

4.2.3.3 Random Forest Model

Machine learning methods are often viewed as black box approaches linking inputs to outputs using a complex set of functions. The main goal of this work is to understand the microstructural attributes causing hotspot formation. Hence we choose a non-black-box machine learning approach: a random forest (RF) algorithm which is described in section 3.1.2. The number of decision trees, the number of features, and the depth of the decision tree are some of the random forest model hyper-parameters. The hyper parameters are chosen using a random grid search by comparing the cross validation performance. The RF algorithm was implemented in Python using the scikit-learn implementation [64]. Note that we tried a number of other tree based models like XGBoost ([101]) and Gradient Boosted trees [102] which did not result in an improvement in the model AUC.

Our dataset was constructed for a set of representative textures as shown in Figure 4.4a. For each texture kind, stochastically different datasets are created via multiple microstructure instantiations. The location of stress hotspots is affected by texture, geometry and the constitutive parameters. For materials consisting of equiaxed microstructures under uniaxial tension, we compare two machine learning frameworks to capture the variation caused by a texture kind:

- **Partition models:** a different RF model is trained for each texture kind and the AUC score is reported using k-fold cross validation for each model (average of validation performance on each microstructure in a texture kind).
- **Mixed-model:** a single RF model is trained on all the microstructures with different textures, and the AUC score is reported using k-fold cross validation (average of validation performance on 2 randomly chosen microstructures from each texture kind).

4.2.3.4 Feature Importance Metrics

Random forest models have an embedded metric known as the permutation accuracy importance (PAI) [62, 65], which can be used to get feature importance scores to understand the contribution of each feature in predicting stress hotspots. However, this metric is prone to correlation bias due to preferential selection of correlated features during the tree building process [103]. The process of choosing the best feature selection algorithm is discussed in Chapter 6, where we arrive at the FeaLect method [104] for feature selection.

To gain data driven insights, we compute feature importances using state of the art FeaLect method [104]. It is a robust feature selection method that computes the feature importance using LASSO (L1) regularization [105]. We first oversample the dataset to balance the population of the two classes. The oversampled dataset was bootstrapped 100 times. In each random subset, various linear models are fitted using the lars method [106], maintaining the regularization strength such that only 10 features are selected by LASSO. Features are scored in each model depending on their tendency to be selected by LASSO in each model. Finally, these relevance-orderings are averaged to give the feature importance on an absolute scale. We used the R implementation of FeaLect to compute our results [107].

4.3 Results and Discussion

The FCC dataset consists of six different textures as shown in Figure 4.4a. Figure 4.4b shows the grain averaged stress distribution in each texture class. The distributions are all left-tailed; both the magnitude and sharpness of the stress peaks vary with texture kind. The optimized RF model hyper-parameters are max_depth=8 and

TABLE 4.1: Cross validation AUCs for FCC materials for mixed and partition models

Texture kind	Partition Model AUC (%)	Mixed Model AUC (%)
1	71.27 ± 1.25	72.07
2	73.91 ± 3.69	73.10
3	66.13 ± 2.74	67.83
4	73.61 ± 6.90	79.84
5	74.01 ± 3.25	76.31
6	72.65 ± 3.07	75.01
All		74.03 ± 3.72

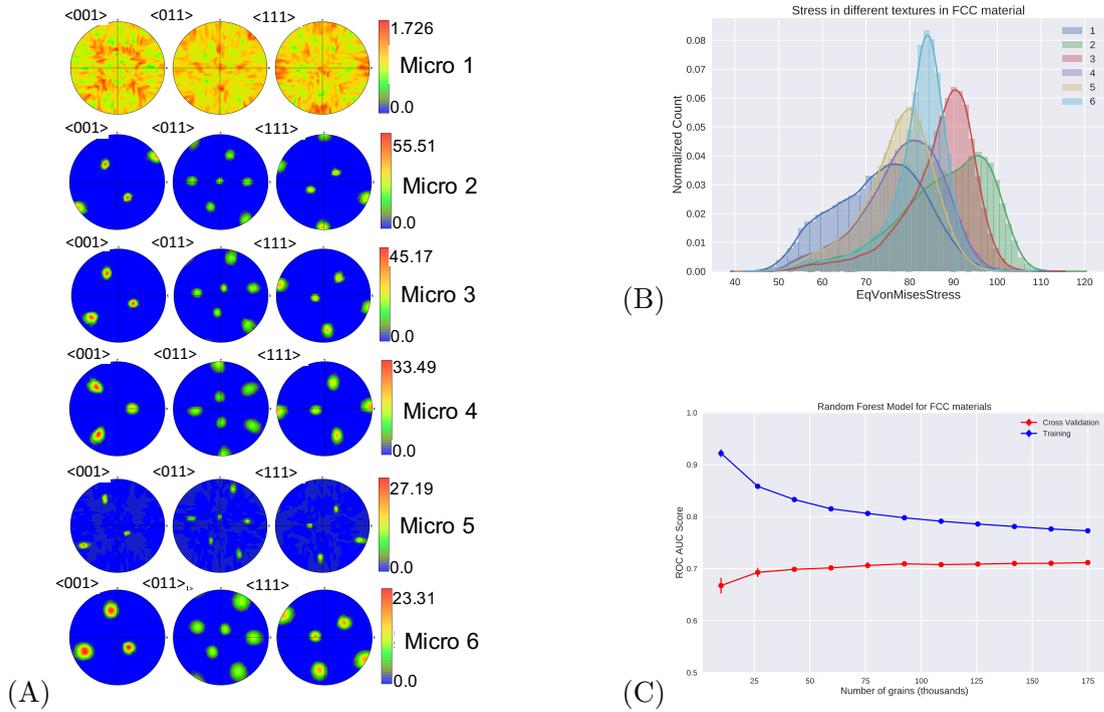


FIGURE 4.4: (a) Representative pole figures for six different FCC textures. (b) Histograms of grain averaged stress in with different textures of FCC materials and (c) Learning curve for random forest model built on all textures of FCC materials

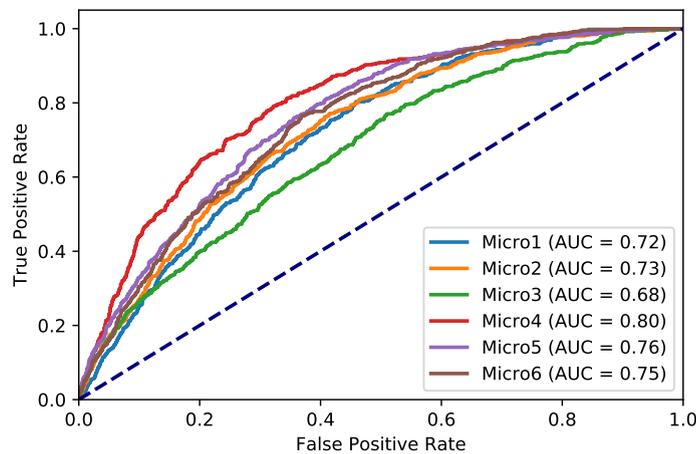


FIGURE 4.5: ROC curves of the mixed microstructure model for validation microstructures in each representative FCC texture. The AUC is the area under this curve.

num_estimators=1200. From table 4.1, it is seen that a single model (Mixed-Model) built on all textures performs better at predicting hotspots than the set of Partition models built on each texture class separately. Not only is the Mixed Model average AUC higher than the average AUC for the Partition Models, but the Mixed Model also yields AUCs equal to or higher than the Partition Models for each individual texture class. The

TABLE 4.2: Pearson Correlation Coefficients between features and Stress hotspots for FCC materials.

Feature	Correlation Coefficient	p-value
001_IPF_0 (Distance between loading direction and $\langle 100 \rangle$)	0.070	0.0
001_IPF_1 (Distance between loading direction and $\langle 110 \rangle$)	-0.144	0.0
001_IPF_2 (Distance between loading direction and $\langle 111 \rangle$)	-0.169	0.0
Schmid Factor	-0.299	0.0
Average Slip Transmission Factor m'	0.038	0.0
Average Misorientation angle	-0.556	0.0
Distance to nearest grain boundary	-0.004	0.13
Distance to nearest triple junction	-0.002	0.49
Distance to nearest quadruple point	-0.002	0.312
Grain Size (NumCells)	-0.003	0.275
FeatureBoundaryElementFrac	0.004	0.16

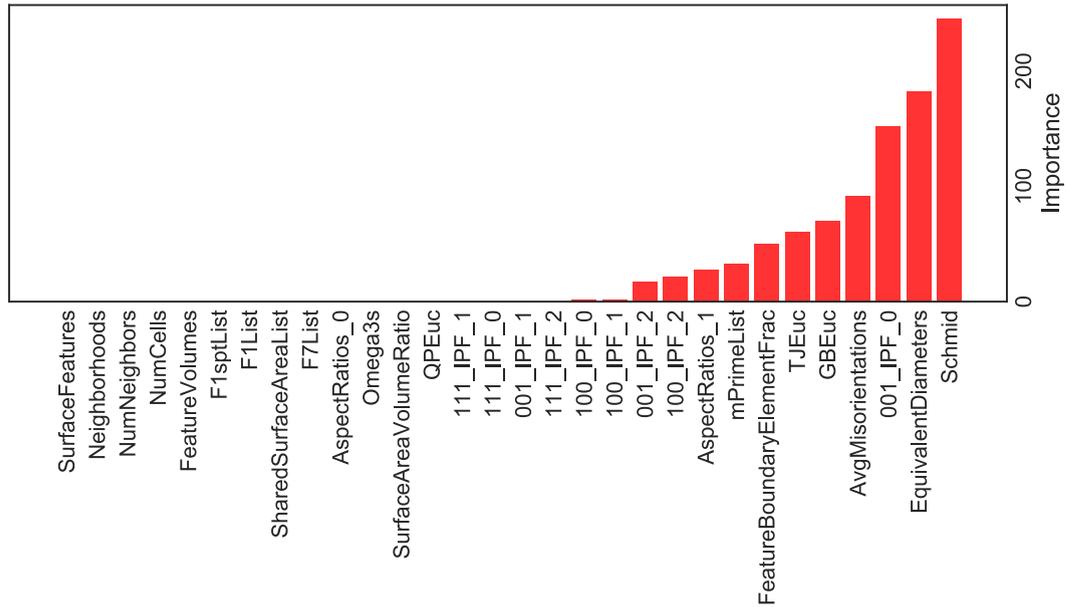


FIGURE 4.6: Variable importance in FCC materials: Both texture and geometry derived microstructural features are selected by FeaLect algorithm. The importance scale is arbitrary and set by Fealect

Mixed Model average AUC = $74.03 \pm 3.72\%$, indicating hotspot prediction well above random chance. This signals the existence of overarching rules causing stress hotspots, independent of the individual textures. We find that adding a categorical feature denoting the texture kind does not result in an improvement in the model performance, so the Mixed Model does not use texture class as a feature.

The ROC curves (described in section 3.2.2) for the mixed microstructure model computed using cross-validation on representative microstructures are shown in Figure 4.5. We can see that the mixed microstructure model performs better on some texture classes, but it performs much better than random chance denoted by the dashed blue line.

From the learning curve for the mixed model shown in Figure 4.4c, since there is a gap between the training and validation scores, we can conclude that the model is suffering from bias, and the model performance can be improved by either increasing the feature space, or by changing the model algorithm. It is difficult to argue whether we have exhausted the space of important crystallographic features, but not many geometric features have been utilized in these models. It is possible that since the model suffers from bias, the *missing features* come from long range geometric features which are not captured in the current feature set.

Figure 4.6 shows the feature importances for the mixed microstructure model calculated using the FeaLect method. The FeaLect plot reveals the set of most important features, which have minimum correlation between them. The Schmid factor of the $\langle 110 \rangle \{111\}$ slip system is the most important feature. We calculated the Pearson correlation [108] between the important features and stress hotspots (table 4.2), and found that hotspots tend to form in grains with lower Schmid factor for FCC materials.

Another important crystallographic feature selected by FeaLect is (Figure 4.6) the distance of the loading direction [001] from the Inverse pole figure corner representing the $\langle 100 \rangle$ crystal direction (001_IPF_0). The elastic modulus for FCC materials is anisotropic; it is highest along $\langle 111 \rangle$ and lowest along $\langle 100 \rangle$ crystal directions. The distance from the 3 corners of the [001] inverse pole figure projection gives an idea about the modulus of hot grains. From table 4.2 we can see that the individual Pearson correlations between the hotspots and these features are very small. We see a positive correlation for the distance of the loading direction from the $\langle 100 \rangle$ corner (001_IPF_0) and negative for the others (001_IPF_1 and 001_IPF_2). Thus the hot grains have a loading direction closer to the $\langle 111 \rangle$ and $\langle 110 \rangle$ corners as compared to the $\langle 100 \rangle$ corner, i.e. hot grains have a higher elastic modulus.

The average misorientation and the slip transmission metric (m-prime) between a grain and its nearest neighbors are the other selected texture derived features. From table 4.2, we see that hot grains have lower average misorientation and a higher slip transmission metric m-prime, i.e. they tend to be grains whose slip systems are more

compatible with their neighbors in agreement with [109]. We take this result with a grain of salt because dislocation movement occurs by a combination of active slip systems and geometrically favorable slip transmission.

Among the geometry derived features, we see that grain size, shape averaged distance from triple junctions and grain boundaries, grain aspect ratio and the fraction of grain lying on the periodic boundary during simulation are important. From table 4.2, we see that hotspots lie closer to grain boundaries, triple junctions and quadruple points i.e. they form in smaller grains. This result is in agreement with [21] where stress hot spots were found to lie closer to microstructural features. Note that although hot spots form in smaller grains, all small grains are not equally probable to stress concentrations. The grain size becomes important in association with crystallography derived features.

It has been shown that the stress field calculated by EVPFFT can spike at the periodic microstructure boundaries, because voxels at the surface are different from the interior [110]. The fraction of a grain lying on the periodic boundary is an important feature in deciding stress hotspots, and captures this simulation artifact by pointing to a numerical issue in the EVPFFT code.

Taken together, these results demonstrate that both crystallographic and geometric features contribute to the formation of stress hotspots. The power of machine learning and ensembling based methods is in discovering a set of microstructural descriptors that cannot be used individually to build a predictive model. The benchmark predictive power (AUC) of the mixed microstructure model is 74.03% compared to around 52% for a model built on the best individual descriptor. As the materials data science field progresses, we can start utilizing these methods to train models to predict the spatial stress field that evolve due to a complex and collective interaction between crystallographic and geometric parameters.

4.4 Conclusions

- Random forest models can predict stress hotspots with 74.03% AUC in FCC materials under uniaxial tension. The performance of a random forest model trained for all textures in a material is comparable or better than models trained separately for each texture which signals the possibility of common factors causing stress hotspots in a material.

- Both texture and geometry derived features contribute to the predictive power of the machine learning model. Grains that become stress hotspots tend to be smaller in size, have high elastic modulus, lower average misorientation and have slip systems more compatible with their neighbors.
- The model performance in all the cases can be increased by adding more descriptive features. The geometry based features discussed describe some aspect of the grain and its nearest neighborhood. Adding long range connectivity based non-crystallographic features might result in an improvement in the model performance.
- The feature importances can delineate the microstructural characteristics with the highest impact. Using these insights, machine learning models can be used to design experiments to develop an understanding about the different feature contributions to the target problem.

4.4.1 Contributions

In this chapter, we have developed a machine learning framework to capture the effects of changing material and texture parameters on stress hotspots. The microstructural features developed in this work can be applied to a range of problems such as prediction of active slip systems for a given texture and loading condition. These methods are also applicable to HEDM obtained datasets. The feature importance plots are a useful way of determining the most important factors when studying a complex problem with many interacting parameters. As the materials science community moves towards a data driven paradigm, it becomes all the more important to examine these techniques.

Chapter 5

Understanding Stress hotspots in Hexagonal Close Packed Materials

The crystal plasticity in hexagonal close packed materials is complex due to the existence of multiple deformation modes. We study the effect of preferred slip systems and microstructural features that reflect local crystallography, geometry, and connectivity on stress hotspot formation in hcp materials under uniaxial tensile stress. Specifically, we considered two cases, one without any preferred slip systems with critically resolved shear stress (CRSS) ratio of 1:1:1, and second with the ratio of 0.1:1:3 for basal: prismatic: pyramidal slip systems. Our random forest based machine learning models achieve an AUC (area under curve) score of 0.82 for Equal CRSS and AUC of 0.81 for Unequal CRSS case. The results show how data driven techniques can be utilized to predict hotspots as well as pinpoint the microstructural features causing stress hotspot formation in polycrystalline microstructures.

5.1 Introduction

In polycrystalline materials, an applied stress is distributed inhomogeneously, resulting in stress concentrations, termed stress hot spots. An important mechanism for ductile fracture in metals and their alloys is by the growth and coalescence of microscopic voids, which nucleate near stress hotspots ([17]). In chapter 4, we concluded that for fcc materials under uniaxial tensile deformation, stress hotspots tend to form near microstructural

features and usually form in textures corresponding to maxima in Taylor factor. The crystalline anisotropy, which determines the "hard" and soft" directions; also plays an important role. The elastic/ plastic behavior of hexagonal close packed (HCP) materials is more complex due to the inherent anisotropy of a non-cubic crystal structure. HCP materials deform plastically by slip on 4 slip systems: basal $\{0001\}[1120]$, prismatic $\{1010\}[112]$, pyramidal $\langle a \rangle \{1101\}[1120]$ and pyramidal $\langle c + a \rangle$, each with different critical resolved shear stress (CRSS) values [111]. The different slip systems in a hcp crystal are shown in figure 5.1. (Deformation twinning also adds to the complexity but has been ignored in this work.) Deformation textures developed in HCP materials vary due to the unique slip and twinning systems that are activated based on the c/a ratio and the CRSS of basal and non basal slip modes.

To understand polycrystal plasticity and texture development in terms of single crystals, the concept of the single crystal yield surface (SCYS) was developed. The SCYS determines the shears that are activated in a grain and depends on the CRSS ratios between deformation modes, as well as the stress state. The SCYS has been analyzed and derived in detail for BCC materials in [25], for FCC materials in [26] and HCP materials in [27]. [28] showed that the SCYS is topologically invariant in certain domains of CRSS ratios, and leads to a simplified analysis of deformation when slip modes harden at different rates. The CRSS ratio is defined with respect to the basal slip resolved shear strength (τ_{basal}) as:

$$CRSSRatio == \frac{\tau_{prismatic}}{\tau_{basal}} : 1 : \frac{\tau_{pyramidal}}{\tau_{basal}} \quad (5.1)$$

where $\tau_{prismatic}$ and $\tau_{pyramidal}$ are the CRSS of prismatic and pyramidal slip systems respectively. Even if the CRSS of a mode is very high, it might be activated to complete the yield surface to achieve the 5 independent slip modes required by the Taylor and Von-Mises criteria, resulting in a highly anisotropic macroscopic response [86, 91]. The situation is worsened by the need to satisfy compatibility and equilibrium conditions between neighboring grains, and results in the material selecting a spatially inhomogeneous solution to accommodate the macroscopic boundary conditions.

Changing the texture of the material will have the same effect of making some slip systems more favorable than others. Hence in order to understand the evolution of stress hotspots, it is necessary to look into a combination of all these variables: texture, grain shape, c/a ratio, CRSS ratios, slip hardening, twinning, temperature and stress

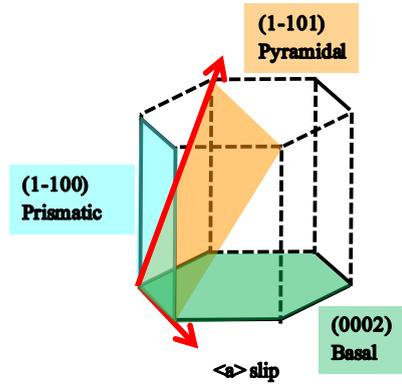


FIGURE 5.1: Schematic of the different slip systems in a hexagonal close packed structure: basal $\{0001\}[1120]$, prismatic $\{1010\}[112]$ and pyramidal $\langle c+a \rangle$. When the tensile axis lies in the $(10\bar{1}1)$ pyramidal plane, the Schmid factor of prismatic $\langle a \rangle$ slip is higher than the basal $\langle a \rangle$ slip.

state. In this work, we keep the temperature constant, and uniaxial tensile deformation is constrained to occur only by 3 slip modes: prismatic, basal and pyramidal $\langle c+a \rangle$ without any twinning or anisotropic slip hardening. The microstructure consists of equiaxed grains and the c/a ratio is fixed. Thus, we can vary the CRSS ratio and crystallographic texture to analyze their impact on stress hotspot formation.

In chapter 4, we used ML methods to analyze stress hotspots in FCC materials. Our model was based on local microstructural features that describe the crystallography (Euler angles, Schmid factor, misorientations) and geometry (grain shape, grain boundary types). The target to was predict whether a grain becomes a stress hotspot based on a feature vector \mathbf{X} whose components are the local microstructural descriptors. In this chapter, we extend this approach to study stress hotspots in HCP materials as a function of texture and compare them among two different HCP materials: an ideal Equal CRSS ratio case where the CRSS ratio is $1 : 1 : 1$ and an Unequal CRSS ratio case of $0.7 : 1 : 3$. We then compare the performance of machine learning models and delineate the microstructural features that contribute the most in predicting stress hotspots.

5.2 Methods

5.2.1 Dataset Generation

We use the Dream.3D package [81] to generate a dataset of synthetic polycrystalline microstructures with a mean grain size of 2.7 microns consisting of ~ 5000 grains each. We study 8 representative textures shown in Figure 5.2. For each representative texture, 9 stochastic microstructure instantiations were created, resulting in ~ 45000 grains per

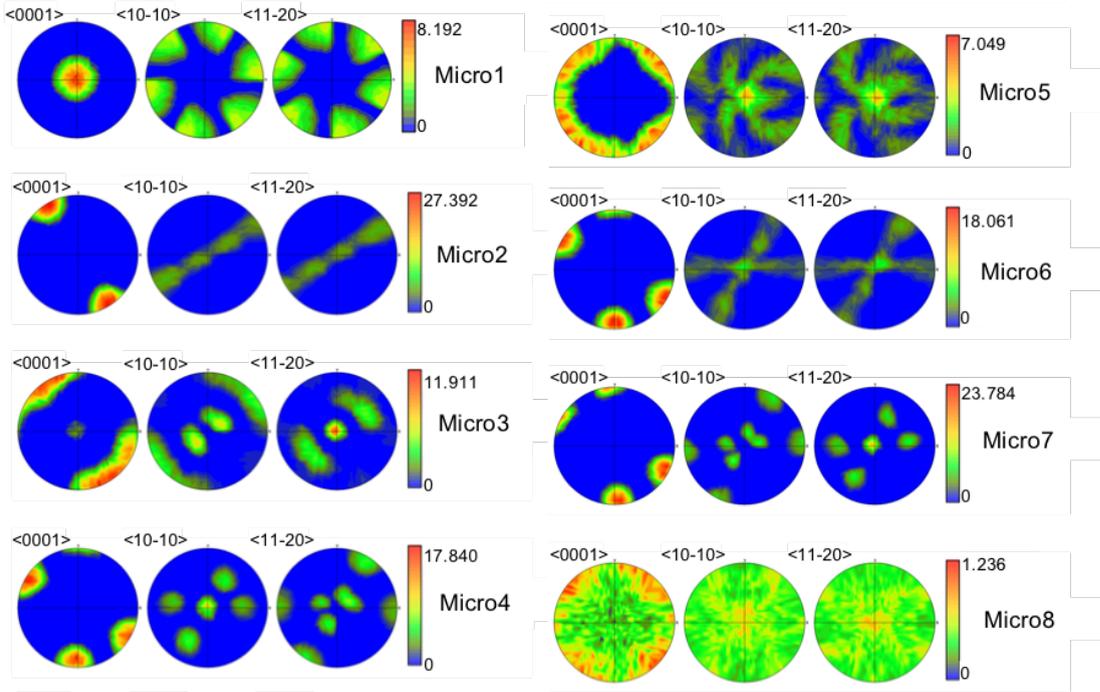


FIGURE 5.2: Representative textures for 8 different HCP textures, the corresponding scale bars show the texture intensity in MRD

texture. The texture intensity for each microstructure instantiation varied from weak (<5 MRD) to strong (>30 MRD), where MRD (multiples of random density) denotes the intensity of a crystallite orientation with respect to a randomly textured material.

The microstructures were then discretized on a $128 \times 128 \times 128$ grid to facilitate the use of EVPFFT (elasto-viscoplastic fast Fourier transform): an image based crystal plasticity formulation [22] to simulate uniaxial tensile deformation. The constitutive model parameters for HCP materials represent a general alpha-titanium alloy and are summarized in appendix B table B.2. The EVPFFT model uses the **Voce hardening** law [45] to model strain hardening as follows:

$$\tau^s(\Gamma) = \tau_0^s + (\tau_1^s + \theta_1^s \Gamma) \left(1 - \exp\left(-\Gamma \left| \frac{\theta_0^s}{\tau_1^s} \right| \right) \right) \quad (5.2)$$

where for a given slip system s , τ_0 is the initial yield stress and θ_0 is the initial hardening rate. $(\tau_0 + \tau_1)$ is the back-extrapolated stress and θ_1 is the asymptotic hardening rate. Γ is the accumulated shear in the grains. The Voce hardening parameters were extracted by fitting the VPSC code [34] generated stress-strain curve to the experimentally obtained curve as shown in the appendix (table B.1).

5.2.2 Microstructural Descriptors

The dataset generated consists of voxelwise representation of the stresses in each microstructure. The spatially resolved stress field is then averaged grain wise to minimize the impact of numerical artifacts and small-scale fluctuations. The resultant Von Mises stress distribution is then thresholded above 90th percentile to designate stress hotspots following the same procedure as in chapter 4 section [82]. This results in 10% of the grains designated as stress hotspots.

Stress distribution in a microstructure is affected by crystallography as well as grain neighborhood and geometry. Hence we develop microstructural features describing the crystallography, geometry and connectivity of grains; and use these as input features to a machine learning algorithm that predicts whether a grain is hot or not. We have developed a number of microstructural features in chapter 4. Along with those features, we include additional HCP material specific features describing the crystallography and geometry to be used in this paper.

The crystallographic descriptors include distance from inverse pole figure corner, which quantifies a grain's orientation with respect to the [001], [010] and [100] directions in the sample frame; features quantifying the misorientation between a grain and its neighbors such as minimum, median, mean etc. and Schmid Factors for each of the basal, prismatic and pyramidal slip systems for each grain. Due to inherent anisotropy in HCP materials, the orientation of the HCP c-axis with respect to the tensile axis is also a good descriptor.

The geometry based descriptors include shape averaged Euclidean distance from special points such as grain boundaries, triple junctions and quadruple points. Features based on grain shape include grain size, equivalent diameter, volume, number of contiguous neighbors, number of neighbors, grain aspect ratio and surface area to volume ratio. Table A.3 lists the acronyms and descriptions of the features used in this work.

We now have datasets for *Equal CRSS* and *Unequal CRSS* consisting of grain-wise labels denoting a stress hotspot, and grain-wise features for each microstructure. Each dataset has 72 microstructures (8 representative texture kinds and 9 microstructures per texture). Although both the Equal and Unequal CRSS ratio data sets represent HCP materials with the same c/a ratio, constitutive parameters are different, so they fundamentally represent different materials. Hence a machine learning model is built

for each case to predict whether a stress hotspot forms in a given grain.

5.2.3 Machine Learning Methods

Since the stress distribution in a microstructure is impacted by a complex interplay of crystallography, geometry and connectivity, we want to build a predictive model which minimizes the assumptions about which features cause hotspot formation. Machine learning models present this opportunity by providing a statistical framework to create connections between the target to be predicted (hotspot) and the features describing it [112].

In this work, we utilize a decision tree based model known as the random forest (RF) algorithm [62, 65] to build our classification model. RF models are very fast and easy to fit, can handle all kinds of features (numerical, categorical) and deal with missing features or data effectively. RF models also provide the importance of the predictor variables [65], which help us gain additional insights into which microstructural descriptors most impact hotspot formation. The model hyper-parameters include the number of decision trees, number of features and the depth of the decision trees. The hyper parameters are chosen using a random grid search by comparing the cross validation performance. The model was implemented using the Scikit-learn library in Python [64]. The RF model has been described in detail in section 3.1.2

The aim is to classify the grains as "hot" or normal, which is an unbalanced binary classification problem. Hence we use the area under the receiving operator characteristic curve (AUC) metric to compare model performance [98]. If the classifier is no better than random guessing, the AUC will be around 0.50. A good classifier has an $AUC \sim 1$. Since varying the texture also has an impact on the location of stress hotspots (chapter 4), we build two kinds of models for each case: **Partition models** and **Mixed-model** as described in chapter 4. We compare the performance of these two kinds of models in HCP materials and contrast it with our result for FCC materials where Mixed-models performed better than Partition models.

Finally, we use the FeaLect method ([104]) to extract feature importances from the dataset, which is then used to derive data driven insights. FeaLect is a state of the art feature selection algorithm that is robust to correlation between the features, and selects the subset of features most highly correlated to the target but least correlated to one another. First the dataset was oversampled to balance the population of the two

TABLE 5.1: Cross validation AUCs (%) for mixed and partition models in equal and Unequal CRSS ratio case of HCP materials

Texture kind	Equal CRSS		Unequal CRSS			
	Partition model AUC	Mixed AUC	model	Partition model AUC	Mixed AUC	model
1	87.47 ± 0.67	87.84		71.87 ± 0.65	71.73	
2	66.20 ± 8.65	77.93		86.61 ± 0.75	85.78	
3	74.89 ± 6.44	90.51		72.52 ± 3.35	75.94	
4	69.89 ± 11.82	78.27		83.20 ± 5.07	82.89	
5	83.22 ± 11.98	89.79		76.78 ± 3.74	73.72	
6	79.89 ± 10.03	86.12		77.62 ± 6.46	87.61	
7	73.39 ± 8.52	64.19		85.31 ± 1.78	85.43	
8	85.48 ± 0.51	85.37		87.76 ± 0.61	86.22	
All	77.55 ± 7.66	82.50 ± 8.22		80.21 ± 6.33	81.18 ± 5.94	

classes. It was then bootstrapped into 100 subsets. In each random subset, linear models are fitted using least angle regression (LARS) method ([106]) with the regularization strength such that only 10 features are selected in each model. Features are scored on their tendency to be selected in each model. Finally, these scores are averaged to give the feature importance on an absolute scale. We used the R implementation of FeaLect to compute our results ([107]).

5.3 Results and Discussion

For the equal CRSS material, the ratio of basal $\langle a \rangle$: prismatic $\langle a \rangle$: pyramidal $\langle c + a \rangle$ CRSS is 1:1:1. It is worth noting that this CRSS ratio is not observed in α -Ti, and represents an ideal HCP material with isotropic slip systems. Figure 5.3a shows the representative grain averaged stress distribution in each texture class for the Equal CRSS ratio case: the stress distributions are all right tailed.

For the Unequal CRSS ratio case, uniaxial tensile deformation is simulated with the same microstructure set as the Equal CRSS ratio case, but using different constitutive parameters. The CRSS ratio chosen is basal $\langle a \rangle$: prismatic $\langle a \rangle$: pyramidal $\langle c + a \rangle = 1 : 0.7 : 3$. This CRSS ratio is selected to better represent α -Ti ([113]). It was observed that due to the inhomogeneity in CRSS values, texture heavily influences the macroscopic response. Figure 5.3b shows the grain averaged stress distribution in each texture class. The stress distributions change character between different textures.

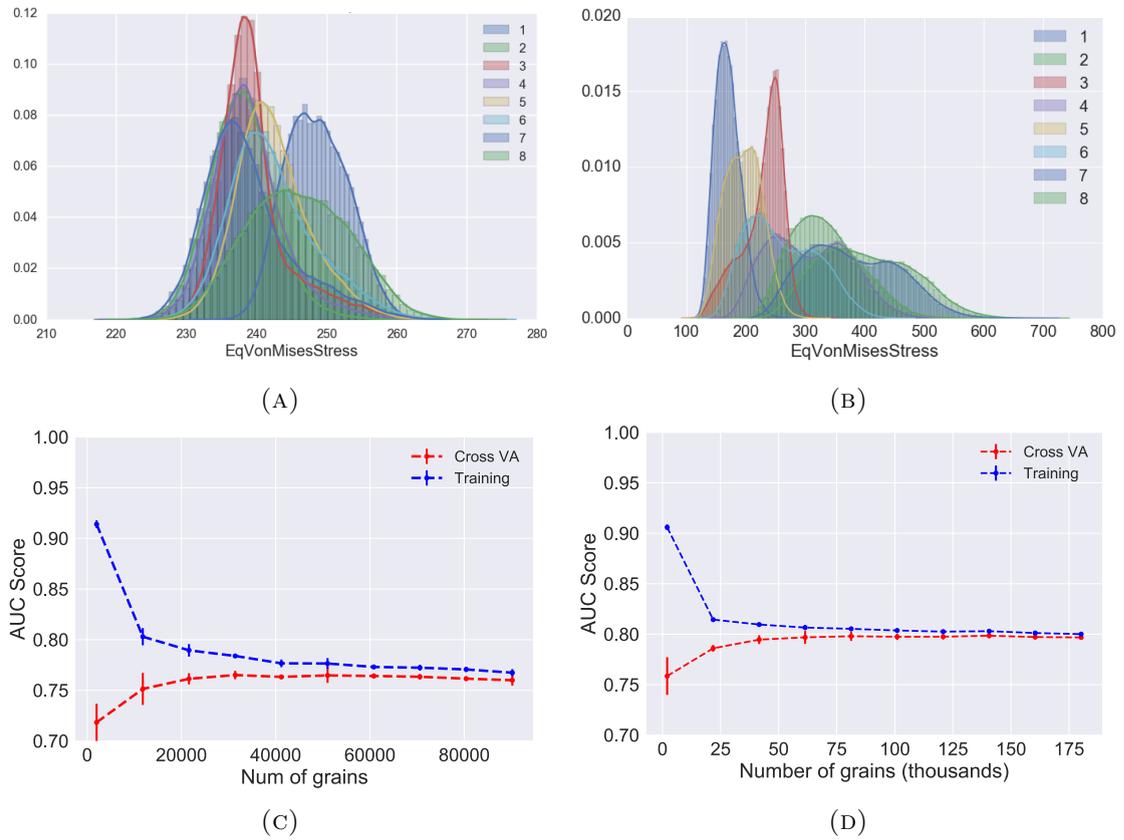


FIGURE 5.3: Histograms of grain averaged stress with different textures in HCP materials with (a) Equal CRSS ratio and (b) Unequal CRSS ratio. The corresponding learning curves for Mixed-Micro model in HCP materials with (c) Equal CRSS ratio and (d) Unequal CRSS ratio

TABLE 5.2: Pearson Correlation Coefficients between features and Stress hotspots for HCP materials

Feature	Equal CRSS		Unequal CRSS	
	Correlation Coefficient	p-value	Correlation Coefficient	p-value
theta	0.002	0.27	-0.0029	0.1083
phi	0.089	0.0	0.1276	0.0
Basal $\langle a \rangle$ Schmid	0.5428	0.0	-0.3933	0.0
Prismatic $\langle a \rangle$ Schmid	-0.5567	0.0	0.490	0.0
Pyramidal $\langle a \rangle$ Schmid	-0.0629	0.0	0.490	0.0
Pyramidal $\langle c + a \rangle$ Schmid	0.1181	0	-0.1777	0.0
GBEuc	-0.0027	0.14	-0.0084	0.0
TJEuc	-0.0024	0.20	-0.0094	0.0
QPEuc	-0.0021	0.25	-0.0052	0.005
Equivalent Diameter	-0.0023	0.22	-0.0087	0.0

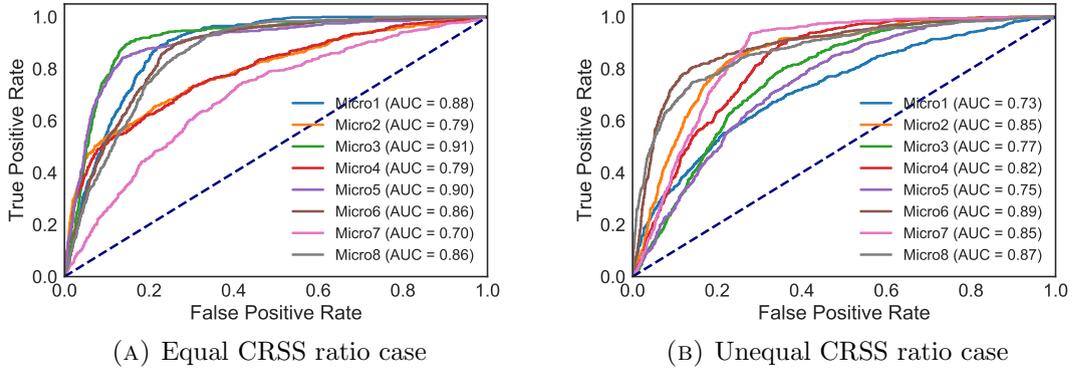


FIGURE 5.4: ROC curves of the mixed microstructure model for validation microstructures in each representative HCP texture for both the CRSS ratio cases. The AUC is the area under this curve.

Partition and Mixed random forest models were computed for both the CRSS ratio cases separately. The optimized model hyper-parameters are: max depth=8 and num estimators=1200. Table 5.1 reports the AUC score using 9 fold cross validation for Partition models and 8 fold cross validation for Mixed models. It is found that Mixed models perform comparably or better than the Partition models for both the datasets. This is a surprising positive result, as it eliminates the need for training different models for each material texture class. The benchmark predictive power (AUC) of the mixed microstructure model is $82.5 \pm 8.22\%$ and $81.18 \pm 5.94\%$ in the Equal and Unequal CRSS ratio cases respectively. The ROC curves (described in section 3.2.2) for both the CRSS ratio cases are shown in Figure 5.4. From Figure 5.4, we see that the RF models perform differently between corresponding representative textures for each CRSS ratio case. This means that it is easier to predict stress hotspots in certain textures. However, note that the texture having better model prediction for Equal CRSS case does not transfer to Unequal CRSS case. For example, Micro3 has an AUC of 0.91 in Equal CRSS whereas it has an AUC of 0.77 in Unequal CRSS case.

The learning curves for the mixed models for the two datasets are shown in figure 5.3c and 5.3d. The training and validation model performance seem to converge in both cases which means the model performance can be improved by either increasing the feature space, or by using a more complex model algorithm.

Figure 5.5 shows the feature importances for the mixed-microstructure model calculated for the Equal CRSS and Unequal CRSS ratio cases using the FeaLect algorithm. In the case of an Equal CRSS ratio (green bars), the basal Schmid factor is the

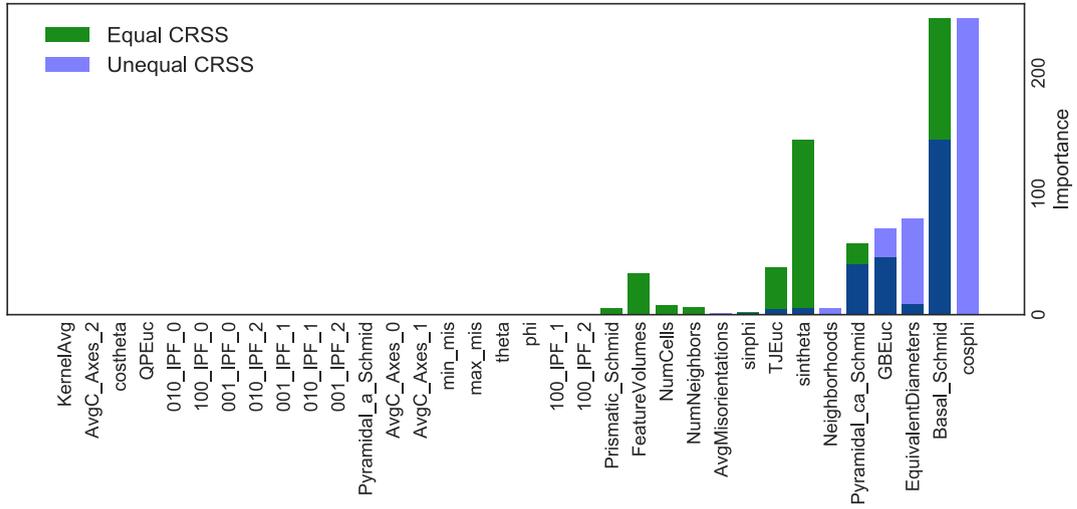


FIGURE 5.5: FeaLect variable importance in HCP materials showing selected features for Equal CRSS (green) and Unequal CRSS (blue)

most important feature, followed by the HCP-c axis orientation ($\sin\theta$) and the pyramidal $\langle a \rangle$ Schmid factor. We calculated the Pearson correlation between the important features and stress hotspots (table 5.2), and found that hotspots tend to form in grains with higher polar and azimuthal angle of the HCP-c axis, which translates to grains with higher basal Schmid factor, which is proportional to the cosine of theta. The elastic modulus for HCP materials (Ti) has an angular behavior which is captured by θ . The elastic modulus is highest along $\langle 0001 \rangle$ direction and lowest in the $[0001]$ plane. Hence hotspots form in grains with lower elastic modulus. This shows the power of feature selection to capture physical effects, since in the absence of heterogeneous slip systems, the stress distribution is impacted by the directionality in elastic modulus which in turn is dependent on theta. This trend is similar to our result in FCC materials from chapter 4; when the material has homogeneous deformation modes, the most important features are those which couple the loading direction and the crystallography. The geometry derived features come next on the feature importance plot, and we found that hotspots lie closer to grain boundaries, triple junctions and quadruple points i.e. form in smaller grains (table 5.2). This result is in agreement with our FCC results (4) and [21] where stress hotspots were found to lie closer to microstructural features.

For the Unequal CRSS ratio material, from figure 5.5 (blue bars), we see that the set of important features are HCP c-axis orientation (phi, theta, basal Schmid factor), grain size, pyramidal $\langle c + a \rangle$ Schmid factor and shape averaged triple junction

distance per grain. The top 3 important features include the grain size (equivalent diameter) in contrast to materials with homogenous deformation modes. From the Pearson correlation coefficients (table 5.2), we observe that hotspots lie closer to grain boundaries, triple junctions and quadruple points, form in grains with low basal and pyramidal $\langle c + a \rangle$ Schmid factor and prefer a high prismatic $\langle a \rangle$ Schmid factor. In the following section, we explore the effect of competing slip systems to better understand the feature importance results.

5.3.1 Role of competing slip systems in stress hotspot formation

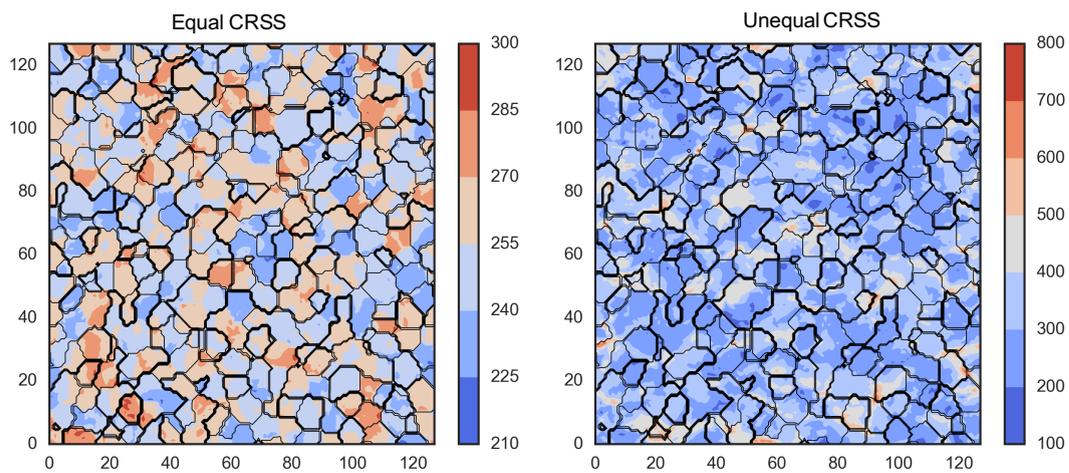
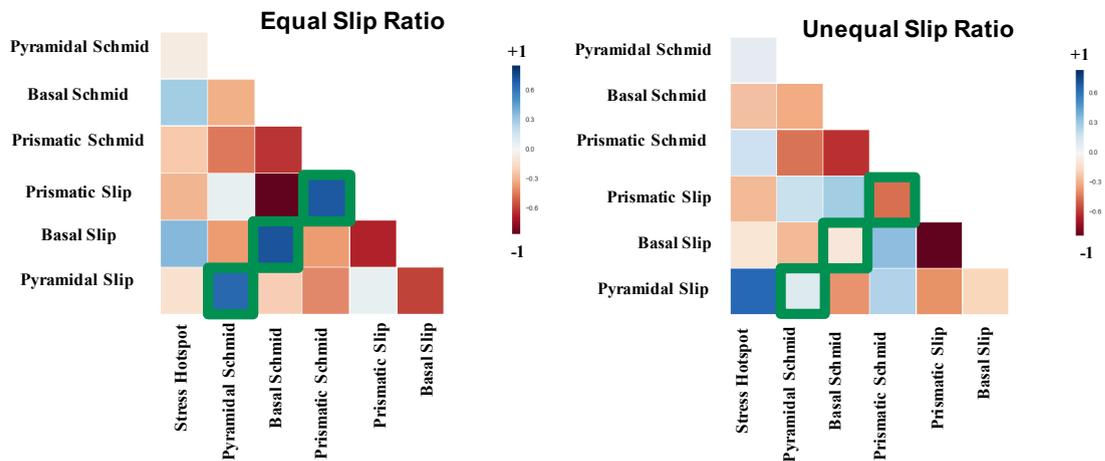


FIGURE 5.6: Cross section of a randomly textured 3-D equiaxed microstructure showing the Von Mises stress distribution under different SCYS topology regimes for a microstructure with random texture.

To compare the effect of the competing slip systems on stress hotspot formation in HCP materials, the set of microstructures with random texture is selected. Figure 5.6 shows the cross section of one of these microstructures, with the spatially resolved Von Mises stress field in the Equal and Unequal CRSS ratio cases. It can be observed that stress hotspots are more pronounced when a limited number of slip systems is available (Unequal CRSS), and for the same microstructure, hotspot location changes with available slip systems. It was found that the skewness of the grain averaged stress histogram for the Equal CRSS case is 0.085 and for the Unequal CRSS case it is an order of magnitude larger, 0.85; that is, when slip systems are limited, a heavy tailed stress distribution is observed. Due to the high CRSS for pyramidal $\langle c + a \rangle$ slip compared to prismatic $\langle a \rangle$ slip, some grains, due to their orientation, are at a disadvantage, because they cannot provide the necessary deformation modes required to close the



(A) Equal CRSS ratio case: The correlation between corresponding slip activities and Schmid factors is highlighted.

(B) Unequal CRSS ratio case: The correlation between corresponding slip activities and Schmid factors is highlighted. There is a strong positive correlation between pyramidal slip fraction and stress hotspot formation.

FIGURE 5.7: Correlation matrix for slip activities and Schmid factors

yield surface. In such grains, the stress climbs very high and there is no clear yield, thus causing the heavy tail.

The grain population was sorted by stress values and divided into 10 bins, each having 10% of the grains. Hence the last bin corresponds to the grains which are stress hotspots. The relative slip activities in each slip system were then compared between these bins, and the mean value of the slip activities in each bin was compared for the Equal and Unequal CRSS cases. We found that the Equal CRSS ratio hotspots have high basal slip fractions, whereas the Unequal CRSS ratio hotspots have higher pyramidal slip fraction. The number of active slip systems was found to be similar in both cases, thus following the single crystal yield surface (SCYS) criterion.

To further understand the slip activities, a correlation matrix between the slip activity and the corresponding Schmid factors is plotted in figure 5.7a and 5.7b. A strong positive correlation between slip activity and the corresponding Schmid factors is observed for the Equal CRSS ratio case, and the correlation is weak for the Unequal CRSS ratio case. The strong correlation in the equal slip case could be due to a more isotropic yield surface. Because the CRSS for all slip systems is the same, many slip systems are activated at the same stress, and the Schmid factor becomes important. In the Unequal CRSS case, the number of available slip systems is smaller, and even if the CRSS of a mode is very high, it might be activated to complete the yield surface to

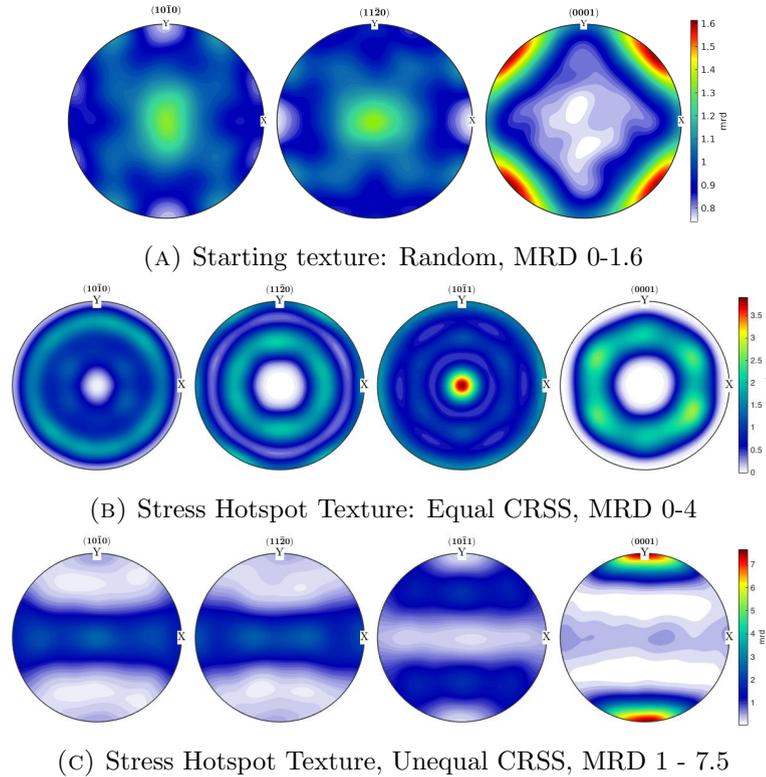


FIGURE 5.8: Pole figures of the initial random texture, hot grains in Equal CRSS case and hot grains in Unequal CRSS case. Note that different scale bars.

achieve 5 independent slip modes.

Figure 5.8 shows the pole figures for the starting microstructure and the hot grains. Starting with a random texture (figure 5.8a), we notice that hotspots form in completely different textures in the two cases.

For the Equal CRSS case, (figure 5.8b), it can be observed that the hotspot loading direction (the z-axis in the sample reference frame) has no preference to align with the $(10\bar{1}0)$ or $(11\bar{2}0)$ planes. However, the loading direction aligns with the $[10\bar{1}1]$ pyramidal pole as seen from the (0001) and $(10\bar{1}1)$ pole figure. When this happens, the loading direction lies in the pyramidal plane as shown in figure 5.1. In this orientation, the Schmid factor favors prismatic slip, so if basal slip is becoming active in these grains, it should mean they have a higher stress.

For the Unequal CRSS case, from the set of hotspot pole figures (figure 5.8c), we can see that there is no preference for the loading direction to align with the prismatic and pyramidal planes. The c-axis aligns with the sample y-axis which means these grains have a low elastic modulus. Since the c-axis is perpendicular to the tensile axis, the deformation along the tensile direction can be accommodated by prismatic slip, and

if pyramidal slip is occurring, it requires a very high stress. From the comparison of pole figures of hot grains, the dominant slip modes cannot be predicted with confidence.

5.4 Conclusions

- Stress hotspots can be predicted with 82.5% AUC in HCP materials with Equal CRSS ratio, and 81.18% AUC in HCP materials with Unequal CRSS ratio using random forest models. We observe that the performance of Mixed-models is comparable to or better than Partition-models. This could mean the existence of common factors independent of the macro-texture which cause stress hotspots in a material.
- A change in material composition will result in altered constitutive parameters, and consequently, the mechanical response. This changes the microstructural descriptors needed, and hence models need to be built for each material.
- Contrasting stress hotspot formation for Equal vs Unequal CRSS ratios in materials with random texture, we observe:
 - Stress hotspots are more pronounced when a limited number of slip systems is available (Unequal CRSS), and for the same microstructure, hotspot location changes with available slip systems.
 - Stress hotspots in the Equal CRSS ratio case have high basal slip fractions and strong positive correlation between slip activity and corresponding Schmid factors, which could be due to an isotropic yield surface.
 - Stress hotspots in the Unequal CRSS ratio case have higher pyramidal slip fraction and weak correlations between corresponding slip activities and Schmid factors. This could be due to the limited number of slip systems.
 - A comparison between the feature importance results reveals that the macro-texture (HCP c-axis orientation) mainly determines stress hotspots in the Equal CRSS ratio case. In the Unequal CRSS ratio case, both crystallography and geometry based features are required to predict stress hotspots.

5.4.1 Contributions

We have successfully demonstrated the applicability of a data driven approach for predicting stress hotspots in different kinds of HCP materials. Using feature importance

plots, we are able to gain objective insights on how hotspot formation varies with material parameters such as CRSS ratio. Hence the framework used in this work is not limited to predicting stress hotspots in HCP materials, but can be extended to various polycrystalline materials, and a wide range of structure-property relationships in materials.

Chapter 6

Feature Selection Techniques

The first step in constructing a machine learning model is defining the features of the data set that can be used for optimal learning. In this work we discuss feature selection methods, which can be used to build better models, as well as achieve model interpretability. We applied these methods in the context of stress hotspot classification problem, to determine what microstructural characteristics can cause stress to build up in certain grains during uniaxial tensile deformation. The results show how some feature selection techniques are biased and demonstrate a preferred technique to get feature rankings for physical interpretations.

6.1 Introduction

Statistical learning methods are gaining popularity in the materials science field, rapidly becoming known as "Materials Data Science". With new data infrastructure platforms like Citrination [4] and the Materials data curation system [5], machine learning (ML) methods are entering the mainstream of materials science. Materials data science and informatics is an emergent field aligned with the goals of the Materials Genome Initiative to reduce the cost and time for materials design, development and deployment. Building and interpreting machine learning models are indispensable parts of the process of curating materials knowledge. ML methods have been used for predicting twinning deformation [80], phase diagrams [114] and guiding experiments and calculations in composition space [10, 115]. Machine learning models are built on learning from "features" or variables that describe the problem. Thus, an important aspect of the machine learning process is to determine which variables most enable data driven insights about the

problem.

Dimensionality reduction techniques (such as principal component analysis(PCA) [116], kernel PCA [117], autoencoders [118], feature compression from information gain theory [119]) have become popular for producing compact feature representations [120]. They are applied to the feature set to get the best feature representation, resulting in a smaller dataset, which speeds up the model construction [121]. Dimensionality reduction has been used by material scientists to establish process-structure-property relationships and for exploratory data analysis to understand trends in a multivariate space [122]. For example, ranking based feature selection methods such as information gain and Pearson correlation have been used during construction of predictive models for fatigue strength of steel [123]. Kalidindi et al. [124] have used 2-point correlations and PCA to describe microstructure-property relationships between local neighborhoods and the localizations in microstructural response. Dey et al. [125] used PCA to analyze the features that cause outliers when predicting bandgaps for new chalcopyrite compounds. Broderick et al. [126] demonstrate how a compact representation (via PCA) makes it easy to visually track the different chemical processing pathways for interpenetrating polymer networks (IPNs) due to changing composition versus changing polymerization. However, dimensionality reduction techniques change the original representation of the features, and hence offer limited interpretability [120]. An alternate method for better models is feature selection. Feature selection is the process of selecting a subset of the original variables such that a model built on data containing only these features has the best performance. Feature selection avoids overfitting, improves model performance by getting rid of redundant features and has the added advantage of keeping the original feature representation, thus offering better interpretability [120].

Feature selection methods have been used extensively in the field of bioinformatics [127], psychiatry [128] and cheminformatics [129]. There are multiple feature selection methods, broadly categorized into Filter, Wrapper and Embedded methods based on their interaction with the predictor during the selection process. The filter methods rank the variables as a preprocessing step, and feature selection is done before choosing the model. In the wrapper approach, nested subsets of variables are tested to select the optimal subset that work best for the model during the learning process. Embedded methods are those which incorporate variable selection in the training algorithm.

We have used random forest models to study stress hotspot classification in FCC (4) and HCP (5) materials. In this paper, we review some feature selection techniques applied to the stress hotspot prediction problem in hexagonal close packed materials, and compare them with respect to future data prediction. We focus on two commonly used techniques from each method: (1) Filter Methods: Correlation based feature selection (CFS) [130], and Pearson Correlation [108]; (2) Wrapper Methods: Fealect [104] and Recursive feature elimination (RFE) [120] and (3) Embedded Methods: Random Forest Permutation accuracy importance (RF-PAI) [65] and Least Absolute Shrinkage and Selection Operator (LASSO) [105]. The main contribution of this article is to raise awareness in the materials data science community about how different feature selection techniques can lead to misguided model interpretations and how to avoid them. We point out some of the inadequacies of popular feature selection methods and finally, we extract data driven insights with better understanding of the methods used.

6.2 Methods

An applied stress is distributed heterogeneously within the grains in a microstructure [20]. Under an applied deformation, some grains are prone to accumulating stress due to their orientation, geometry and placement with respect to the neighboring grains. These regions of high stress, so called stress hotspots, are related to void nucleation under ductile fracture [17]. Stress hotspot formation has been studied in face centered cubic (FCC) [131] and hexagonal close packed (HCP) [132] materials using a machine learning approach. A set of microstructural descriptors was designed to be used as features in a random forest model for predicting stress hotspots. To achieve data driven insights into the problem, it is essential to rank the microstructural descriptors (features). In this paper, we review different feature selection techniques applied to the stress hotspot classification problem in HCP materials, which have a complex plasticity landscape due to anisotropic slip system activity.

Let (x_i, y_i) , for $i = 1, \dots, N$ be N independent identically distributed (i.i.d.) observations of a p -dimensional vector of grain features $x_i \in R^p$, and the response variable $y_i \in 0, 1$ denotes the truth value of a grain being a stress hotspot. The input matrix is denoted by $X = (x_1, \dots, x_N) \in R^{N \times p}$, and $y \in [0, 1]^N$ is the binary outcome. We will use small letters to refer to the samples x_1, \dots, x_N and capital letters to refer to the features X_1, \dots, X_p of the input matrix X . Feature importance refers to metrics used

by various feature selection methods to rank, such as feature weights in linear models or variable importance in random forest models.

6.2.1 Dataset Studied

We use the Unequal CRSS Ratio dataset generated in chapter 4 representing a Titanium like HCP material with CRSS ratio of basal $\langle a \rangle$: prismatic $\langle a \rangle$: pyramidal $\langle c+a \rangle$ = 1 : 0.7 : 3. This dataset contains grain-wise values for equivalent Von Mises stress, and the corresponding Euler angles and grain connectivity parameters.

The grains having stress greater than the 90th percentile of the stress distribution were designated as stress hotspots, a binary target. Thirty four variables to be used as features in machine learning were developed. These features (X) describe the grain texture and geometry and have been summarized in table A.3. We note that these features are not a complete set, and there are long range effects causing stress hotspots. We have taken the first order microstructural descriptors to build stress hotspot prediction models and understand that these models can be improved upon by adding the missing features.

The microstructures contained in this dataset represent eight different kinds of textures, and we validate the machine learning models by leave one texture out validation. This divides the dataset into $\sim 85\%$ training and $\sim 15\%$ validation. Note that since only 10% of the grains are stress hotspots, this is an imbalanced classification problem. Hence, the model performance is measured by the AUC (area under curve), a metric for binary classification which is insensitive to imbalance in the classes. An AUC of 100% denotes perfect classification and 50% denotes no better than random guessing [98].

We first build a decision tree based random forest model [65] for stress hotspot classification using all the thirty four variables. We then rank and select the variables using different feature selection techniques. The selected variables are then used to build random forest models and we observe the improvement in model performance (as described in chapter 4), as well as understand the physics behind stress hotspot formation. The feature rankings are then used to gain insights about the physics behind stress hotspot formation.

6.2.2 Feature Selection Methods

6.2.2.1 Filter Methods

Filter methods are based on preprocessing the dataset to extract the features X_1, \dots, X_p that most impact the target Y . Some of these methods are:

Pearson Correlation [108]: This method provides a straightforward way for filtering features according to their correlation coefficient. The Pearson correlation coefficient between a feature X_i and the target Y is:

$$\rho_i = \frac{\text{cov}(X_i, Y)}{\sigma(X_i)\sigma_Y}$$

where $\text{cov}(X_i, Y)$ is the covariance, σ is the standard deviation [108]. It ranges between $(-1, 1)$ from negative to positive correlation, and can be used for binary classification and regression problems. It is a quick metric using which the features are ranked in order of the absolute correlation coefficient to the target.

Correlation based feature selection (CFS) [130]: CFS was developed to select a subset of features with high correlation to the target and low intercorrelation among themselves, thus reducing redundancy and selecting a diverse feature set. CFS gives a heuristic merit over a feature subset instead of individual features. It uses symmetrical uncertainty correlation coefficient given by:

$$r(X, Y) = 2.0 \times \frac{IG(X|Y)}{H(X) + H(Y)}$$

where $IG(X|Y)$ is the information gain of feature X for the class attribute Y . $H(X)$ is the entropy of variable X . The following merit metric was used to rank each subset S containing k features:

$$\text{Merit}_S = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}}$$

where $\overline{r_{cf}}$ is the mean symmetrical uncertainty correlation between the feature ($f \in S$) and the target, and $\overline{r_{ff}}$ is the average feature-feature inter-correlation. To account for the high computational complexity of evaluating all possible feature subsets, CFS is often combined with search strategies such as forward selection, backward elimination and bi-directional search. In this work we have used the scikit-learn implementation of CFS [133] which uses symmetrical uncertainty [130] as the correlation metric and explores the subset space using best first search [134], stopping when it encounters five

consecutive fully expanded non-improving subsets.

6.2.2.2 Embedded Methods

These methods are popular because they perform feature selection while constructing the classifier, removing the preprocessing feature selection step. Some popular algorithms are support vector machines (SVM) using recursive feature elimination (RFE) [135], random forests (RF) [65] and Least absolute shrinkage and selection operator (LASSO)[105]. We compare LASSO and RF methods for feature selection on the stress hotspot dataset.

Least Absolute Shrinkage and Selection Operator (LASSO) [105]:

LASSO is linear regression with L_1 regularization [105]. A linear model \mathcal{L} is constructed

$$\mathcal{L} : \min_{w \in \mathbb{R}^p} \sum_{i=1}^N \frac{1}{2N} \|y_i - w^T \cdot x_i\|_2^2 + \lambda \|w\|_1$$

on the training data (x_i, y_i) , $i = 1, \dots, N$, where w is a p dimensional vector of weights corresponding to each feature dimension p . The L_1 regularization term ($\lambda \|w\|_1$) helps in feature selection by pushing the weights of correlated features to zero, thus preventing overfitting and improving model performance. Model interpretation is possible by ranking the features according to the LASSO feature weights. However, it has been shown that for a given regularization strength λ , if the features have redundancy, inconsistent subsets can be selected [136]. Nonetheless, Lasso has been shown to provide good prediction accuracy by reducing model variance without substantially increasing the bias while providing better model interpretability. We used the scikit-learn implementation to compute our results [64].

Random Forest Permutation Accuracy importance (RF PAI) [65]:

The random forest is a non linear multivariate model built on an ensemble of decision trees. It can be used to determine feature importance using the inbuilt feature importance measure [65]. For each of the trees in the model, a feature node is randomly replaced with another feature node while keeping all others nodes unchanged. The resulting model will have a lower performance if the feature is important. When the permuted variable X_j , together with the remaining unchanged variables, is used to predict the response, the number of observations classified correctly decreases substantially, if the original variable X_j was associated with the response. Thus, a reasonable measure for feature importance is the difference in prediction accuracy before and after

permuting X_j . The feature importance calculated this way is known as Permutation Accuracy Importance (PAI) and was computed using the scikit-learn package in Python [64].

6.2.2.3 Wrapper Methods

Wrapper methods test feature subsets using a model hypothesis. Wrapper methods can detect feature dependencies i.e. features that become importance in presence of each other. They are computationally expensive, hence often use greedy search strategies (forward selection and backward elimination [137]) which are fast and avoid overfitting to get the best nested subset of features.

Fealect Algorithm [104]: The number of features selected by Lasso depends on the regularization parameter λ , and in the presence of highly correlated features, LASSO arbitrarily selects one feature from a group of correlated features [138]. The set of possible solutions for all LASSO regularization strengths is given by the regularization path, which can be recovered computationally efficiently using the Least Angles Regression (LARS) algorithm [106]. It was shown that LASSO selects the the relevant variables with a probability one and all other with a positive probability [136]. An improvement in LASSO, the Bolasso feature selection algorithm was developed based on this property [136] in 2008. In this method, the dataset is bootstrapped, and a LASSO model with a fixed regularization strength λ is fit to each subset. Finally, the intersection of the LASSO selected features in each subset is chosen to get a consistent feature subset.

In 2013, the FeaLect algorithm, an improvement over the Bolasso algorithm, was developed based on the combinatorial analysis of regression coefficients estimated using LARS [104]. FeaLect considers the full regularization path, and computes the feature importance using a combinatorial scoring method, as opposed to simply taking the intersection with Bolasso. The FeaLect scoring scheme measures the quality of each feature in each bootstrapped sample, and averages them to select the most relevant features, providing a robust feature selection method. We used the R implementation of FeaLect to compute our results [107].

Recursive Feature Elimination (RFE) [135]: A number of common ML techniques (such as linear regression, support vector machines (SVM), decision trees, Naive Bayes, perceptron, e.t.c) provide feature weights that consider multivariate interacting effects between features [120]. To interpret the relative importance of the variables

from these model feature weights, RFE was introduced in the context of support vector machines (SVM) [135] for getting compact gene subsets from DNA-microarray data.

To find the best feature subset, instead of doing an exhaustive search over all feature combinations, RFE uses a greedy approach, which has been shown to reduce the effect of correlation bias in variable importance measures [139]. RFE uses backward elimination by taking the given model (SVM, random forests, linear regression etc.) and discarding the worst feature (by absolute classifier weight or feature ranking), and repeating the process over increasingly smaller feature subsets until the best model hypothesis is achieved. The weights of this optimal model are used to rank features. Although this feature ranking might not be the optimal ranking for individual features, it is often used as a variable importance measure [139]. We used the scikit-learn implementation of RFE with random forest classifier to come up with a feature ranking for our dataset.

6.3 Results and Discussion

Table 6.1 shows the feature importances calculated using filter based methods: Pearson correlation and CFS; embedded methods: Random Forest (RF), Linear regression, Ridge regression (L_2 regularization) and LASSO regression and finally wrapper methods: RFE and Fealect . The shaded cells denote the features that were finally selected to build RF models and their corresponding performances are noted. The input data was scaled by minimum and maximum values to [0,1]. Figure 6.1 shows the correlation matrix for the features and the target.

Pearson correlation can be used for feature selection, resulting in a good model. However, this measure has implicit orthogonality assumptions between variables, and the coefficient does not take mutual information between features into account. Additionally, this method only looks for linear correlations which might not capture many physical phenomenon.

The feature subset selected by CFS contains features with higher class correlation and lower redundancy, which translate to a good predictive model. Although we know grain geometry and neighborhood are important to hotspot formation, CFS does not select any geometry based features and fails to provide an individual feature ranking.

TABLE 6.1: Variable Importance Measures using different methods for HCP materials with Unequal CRSS. The gray shaded cells denote the features selected by the corresponding technique. The features describing grain geometry are shaded in green.

Features	Pearson	CFS	RFE	Random	Regression			Fealect
				Forest	Linear	Ridge	LASSO	
<i>cos</i> ϕ	-0.29	1	1	53.43	27.37	27.36	26.01	245.0
<i>Schmid</i> ₁	-0.39	0	1	0.15	22.72	22.69	14.78	145.00
EquivalentDiameters	-0.01	0	1	0.05	0.15	0.15	0.08	79.47
GBEuc	-0.01	0	1	0.12	0.22	0.22	0.12	71.47
<i>Schmid</i> ₄	-0.18	0	1	0.31	7.29	7.31	10.35	41.27
Neighborhoods	-0.01	0	22	0.01	0.10	0.10	0.00	5.53
<i>sin</i> θ	0.48	1	1	8.74	74.78	74.61	52.99	5.00
TJEuc	-0.01	0	2	0.07	0.97	0.97	0.44	4.93
<i>sin</i> ϕ	0.14	1	16	0.03	80.46	79.96	19.17	1.0
AvgMisorientations	0.31	0	1	8.95	32.08	32.09	32.05	0.83
NumNeighbors	-0.01	0	23	0.01	0.18	0.17	0.03	0.50
<i>Schmid</i> ₃	0.12	0	9	0.03	4.05	4.04	0.00	0.0
<i>Min</i> _{mis}	0.09	0	1	0.72	3.46	3.46	2.19	0.0
<i>AvgC</i> _{Axes} ₁	0.00	0	1	0.22	0.09	0.09	0.00	0.0
<i>Max</i> _{mis}	0.17	0	4	0.02	0.86	0.86	0.03	0.0
NumCells	-0.01	0	18	0.04	1.3e6	0.11	0.21	0.0
<i>Schmid</i> ₂	0.49	0	1	26.80	38.03	37.83	8.37	0.0
KernelAvg	-0.01	0	25	0.0	0.22	0.22	0.00	0.0
010_IPF_1	-0.07	0	5	0.01	0.49	0.49	0.00	0.0
ϕ	0.13	1	3	3.4	66.42	65.94	7.68	0.0
001_IPF_0	0.00	0	11	0.03	0.58	0.57	0.00	0.0
001_IPF_2	0.09	0	21	0.01	0.21	0.24	0.19	0.0
010_IPF_0	0.00	0	12	0.01	0.76	0.76	0.23	0.0
100_IPF_0	0.00	0	10	0.01	0.13	0.13	0.00	0.0
001_IPF_1	0.16	0	15	0.01	0.17	0.14	0	0.0
100_IPF_1	0.07	0	14	0.02	1.10	1.10	0.00	0.0
QPEuc	-0.01	0	6	0.02	0.57	0.57	0.00	0.0
<i>AvgC</i> _{Axes} ₀	0.00	0	7	0.03	0.34	0.34	0.05	0.0
θ	0.00	1	24	0.02	0.04	0.04	0.00	0.0
FeatureVolumes	-0.01	0	13	0.04	1.3e6	0.11	0.00	0.0
010_IPF_2	-0.04	0	17	0.01	0.79	0.79	0.00	0.0
<i>AvgC</i> _{Axes} ₂	0.00	0	8	0.01	2.9e4	0.07	0.00	0.0
100_IPF_2	0.04	0	19	0.01	1.21	1.20	0.00	0.0
<i>cos</i> θ	0.00	1	20	0.01	2.9e4	0.07	0.00	0.0

Random Forest model AUC without feature selection: 71.94%

Random Forest model AUC with selected features (%)								
training	84.02	82.51	84.24	83.82	84.20	84.19	84.31	84.28
validation	80.46	80.45	80.73	80.19	80.72	80.61	80.83	80.75

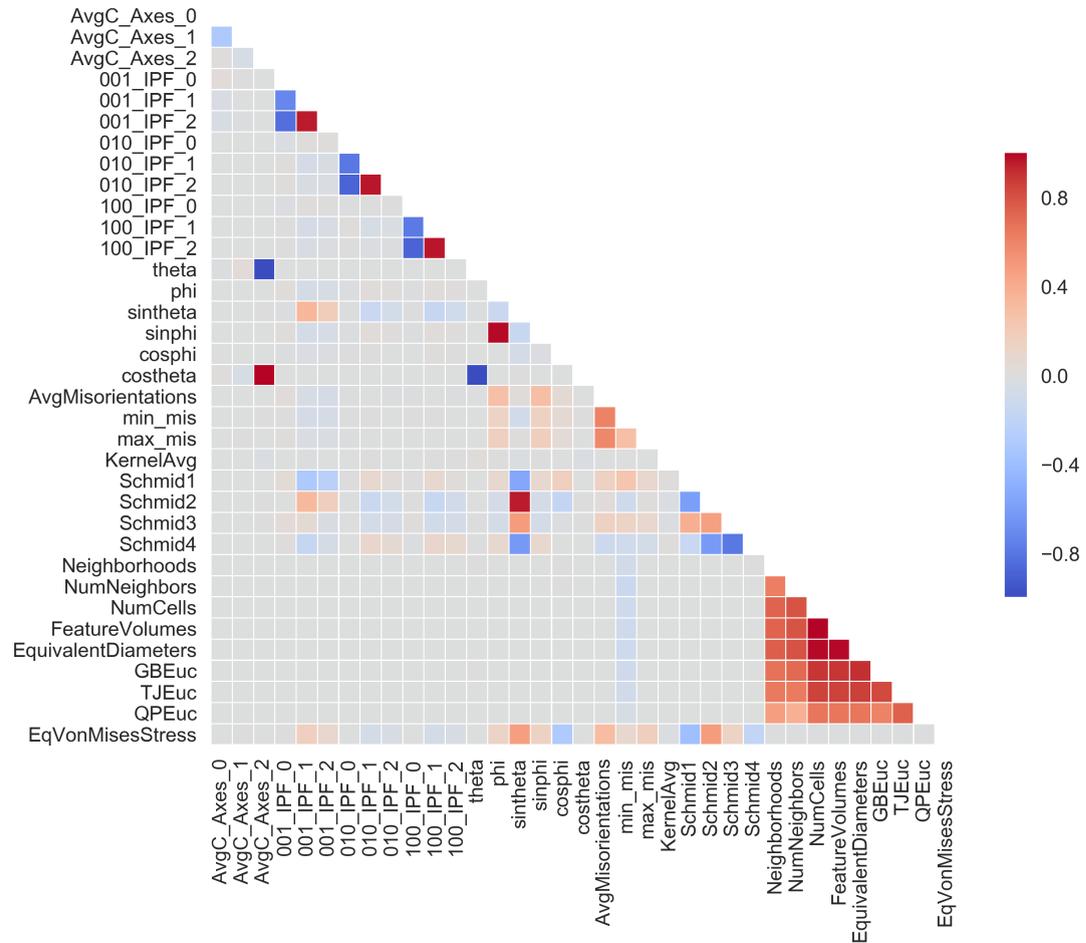


FIGURE 6.1: Pearson Correlation matrix between the target (EqVonMisesStress) and all the features

Linear regression, ridge regression and Lasso are highly correlated linear models. A simple linear model results in huge weights for some features (NumCells, FeatureVolumes), likely due to overfitting, and hence is unsuitable for deducing variable importance. Ridge regression compensates for this problem by using L_1 regularization, but the weights are distributed among the redundant features, which might lead to incorrect conclusions. LASSO regression overcomes this problem by pushing the weights of correlated features to zero, resulting in a good feature subset. The top five ranked features by LASSO with regularization strength of $\lambda = 0.3$ are : $\sin\theta$, AvgMisorientations, $\cos\phi$, $\sin\phi$ and Schmid.1. The first geometry based feature ranks 10th on the list, which seems to underestimate the physical importance of such features. A drawback of deriving insights from LASSO selected features is that it arbitrarily selects a few representatives from the correlated features, and the number of features selected depends heavily on the regularization strength. Thus the models become unstable, because changes in training

subset can result in different selected features. Hence these methods are not ideal for deriving physical insights from the model.

Random forest models also provide an embedded feature ranking module. The RF-PAI importance seems to focus only on the hcp 'c' axis orientation derived features ($\cos\phi, \sin\theta$), average misorientation and the Prismatic $\langle a \rangle$ Schmid factor, while discounting most of the geometry derived features. RF-PAI suffers from correlation bias due to preferential selection of correlated features during tree building process [103]. As the number of correlated variables increases, the feature importance score for each variable decreases. Often times the less relevant variables replace the predictive ones (due to correlation) and thus receive undeserved, boosted importance [140]. Random forest variable importance can also be biased in situations where the features vary in their scale of measurement or number of categories, because the underlying Gini gain splitting criterion is a biased estimator and can be affected by multiple testing effects [141]. From Figure 6.1, we found that all the geometry based features are highly correlated to each other, therefore deducing physical insights from this ranking is unsuitable.

Hence, we move to Wrapper based methods for feature importance. Recursive feature elimination (RFE) has been shown to reduce the effect of the correlation on the importance measure [139]. RFE with underlying random forest model selects a feature subset consisting of two geometry based features (GBEuc and EquivalentDiameter), however, it fails to give an individual ranking among the features.

FeaLect provides a robust feature selection method by compensating for the uncertainty in LASSO due to arbitrary selection among correlated variables, and the number of selected variables due to change in regularization strength. Table 6.1 lists the Fealect selected variables in decreasing order. We find that the top two important features are derived from the grain crystallography, and geometry derived features come next. This suggests that both texture and geometry based features are important. Using linear regression based methods such as these tell us which features are important by themselves, as opposed to RF-PAI which indicates the features that become important due to interactions between them (via RF models) [120]. The Fealect method provides the best estimate of the feature importance ranking which can then be used to extract physical insights. This method also divides the features into 3 classes: informative, irrelevant features that cause model overfitting and redundant features [104]. The most informative features are: $\cos\phi$, Schmid_1, EquivalentDiameter, GBEuc, Schmid_4,

Neighborhoods, $\sin\theta$ and TJEuc. The irrelevant features are $\sin\phi$ and AvgMisorientations (which cause model overfitting). The remaining features are redundant.

A number of selected features directly or indirectly represent the HCP c-axis orientation, such as $\cos\phi$, $\sin\theta$ and basal Schmid factor (*Schmid_1*), which is proportional to $\cos\theta$. It is interesting that pyramidal $\langle c + a \rangle$ Schmid factor (*Schmid_4*) is chosen as important. From Figure 6.1, we can see that hot grains form where θ, ϕ maximize $\sin\theta$ and $\sin\phi$ i.e. $\theta \sim 90, \phi \sim 90$. This means that the HCP c-axis orientation of hot grains aligns with the sample Y axis, which means these grains have a low elastic modulus. Since the c-axis is perpendicular to the tensile axis (sample Z); the deformation along the tensile direction can be accommodated by prismatic slip in these grains, and if pyramidal slip is occurring, it means they have a very high stress [132]. This explains the high importance of the pyramidal $\langle c + a \rangle$ Schmid factor. From the Pearson correlation coefficients in Figure 6.1, we can observe that the stress hotspots form in grains with low basal and pyramidal $\langle c + a \rangle$ Schmid factor, high prismatic $\langle a \rangle$ Schmid factor, and higher values of $\sin\theta$ and $\sin\phi$.

From Figure 6.1, we can see that all the grain geometry descriptors do not have a direct correlation with stress, but are still selected by Fealect. This points to the fact that these variables become important in association with others. We analyzed these features in detail in [132] and found that the hotspots lie closer to grain boundaries (GBEuc), triple junctions (TJEuc), and quadruple points (QPEuc), and prefer to form in smaller grains.

There is a subtle distinction between the physical impact of a variable on the target vs. the variables that work best for a given model. From table 6.1, we can see that a random forest model built on the entire feature set without feature selection has an AUC of 71.94%. All the feature selection techniques result in an improvement in the performance of the random forest model to a validation AUC of about 81%. However, to draw physical interpretations, it is important to use a feature selection technique which: 1) keeps the original representation of the features, 2) is not biased by correlations/redundancies among features, 3) is insensitive to the scale of variable values, 4) is stable to the changes in the training dataset, 5) takes multivariate dependencies between the features into account, and 6) provides an individual feature ranking measure.

6.4 Conclusions

In this work, we have surveyed different feature selection techniques by applying them to stress hotspot classification problem. These techniques can be divided into three categories: filter, embedded and wrapper. We have explored the most commonly used techniques under each category. It was found that all the techniques lead to an improvement in the model performance, and are suitable for feature selection to build a better model. However, when the aim is to interpret the model, and understanding which features might be more causal than others, it is essential to note the limitations of different techniques. We found that in the presence of correlated features, the FeaLect method helped us to determine the underlying importance of the features. We find that:

- All feature selection techniques result in $\sim 9\%$ improvement in the AUC metric for stress hotspot classification.
- Correlation based feature selection and Recursive feature elimination are computationally expensive to run, and give only a feature subset ranking.
- Random forest embedded feature ranking is biased against correlated features and hence should not be used to derive physical insights.
- Linear regression based feature selection techniques can objectively denote the most important features, however have their flaws. These methods can be affected by the scale of features, correlation between them, and the dataset itself.
- The Fealect algorithm can compensate for the variability in LASSO regression, providing a robust feature ranking that can be used to derive insights.
- Stress hotspots formation under uniaxial tensile deformation is determined by a combination of crystallographic and geometric microstructural descriptors.
- It is essential to choose a feature selection method that can find this dependence even when features are redundant or correlated.

Chapter 7

Predicting Material Failure Using Deep Learning

This chapter focuses on integrating crystal plasticity based deformation models [22] and 3D convolutional neural networks (CNNs) to divide the grains in a microstructure based on the stress buildup formed during a uniaxial tensile deformation. We use the dataset developed in Chapter 5 for HCP materials with Unequal CRSS ratio to demonstrate how deep learning can be applied to such problems and comment on the future applications of this approach.

7.1 Introduction

In this thesis, we have predicted stress hotspots using random forest models from microstructural descriptors engineered based on material science knowledge. These descriptors are derived from the grain orientations and geometry and do not contain any long range information about the microstructure, but in polycrystals, long range interactions between individual grains and grain boundaries impact microstructural properties such as abnormal grain growth [142], intergranular crack propagation [143], [88], intergranular corrosion [144] and liquid metal embrittlement[145]. In appendix C, through the use of power laws for neighborhood graphs [146], [147], we have demonstrated that stress hotspots have a different neighborhood compared to the normal grains. We also found that the second nearest neighbors of a hot grain are more probable to be hot themselves, which hints at some kind of medium range order between stress hotspots[148]. We also

used graph kernels and spectral methods to extract features from the transgranular network and found that they were important in determining stress hotspots. However all the spectral and graph kernel features we developed could describe the neighborhoods containing only the nearest neighbors, due to the computational complexity associated with the densely connected microstructure network, which prevented us from engineering truly long range descriptors.

Convolutional neural networks (CNNs) have enjoyed recent success in image recognition problems in domains spanning computer vision [149], [66], [150], [151], [152], medical image segmentation [70], [68], video analysis [153] and natural language processing [154], [155], [156]. Because these networks are invariant to small input variations and geometric transformations [157], CNNs have celebrated success in tasks related to 2D images, video, text, speech and 3D volumetric data. A typical CNN architecture contains alternating sets of convolutional layers with non-linearity and pooling layers, which can represent multiscale features of the image thus providing automatic feature extraction. This architecture has been shown to work for data that has a local structure such as images, where pixels that are spatially or temporally nearby are highly correlated [157]. Figure 7.1 shows a typical CNN architecture consisting of an input layer, multiple hidden convolutional layers and an output unit. A convolutional layer extracts the local features at each stage by restricting the receptive field of the kernel used for convolution. The inputs and outputs of convolutional layers are known as feature maps. In each convolutional layer, a kernel is rastered over the input image to get the feature map which contains information about the local features. This map is fed to a nonlinearity such as a rectified linear unit (ReLU) [158] and passed to a pooling layer which subsamples adjacent rectangular regions (cuboidal regions for 3D data) in the feature maps to single values, bringing down the feature map size. This reduces the total number of parameters in the CNN and prevents overfitting[157]. The convolutional layers at successive stages have a larger receptive field, because a small area in a pooled feature map corresponds to a large area in the input image, hence they will capture high-level global context. The final learning task is performed using fully connected layers stacked on top. The network is then trained end to end, which helps in learning hierarchical features of an image [159], [160].

Notice that the stress hotspot classification problem is similar to the semantic image segmentation problem from the field of computer vision. We have 3D microstructures with voxelwise microstructural descriptors containing information about the grain

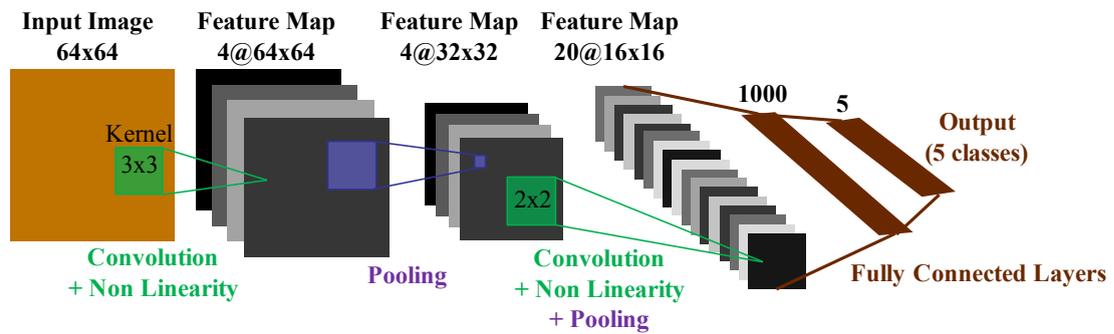


FIGURE 7.1: A typical CNN architecture showing an input image, 2 sets of convolutional, nonlinearity and pooling layers and a fully connected multilayer perceptron at the end to classify the input image into 5 classes.

orientation (Euler angles) and grain size, and we want to segment this microstructure to find grains where stress hotspots form. Hence CNNs are an obvious choice for extracting the multiscale features of a microstructure, based on the texture and geometry information. In this chapter, we develop CNN architectures for the dataset developed in Chapter 5 for HCP materials with Unequal CRSS ratio and comment on the applications and future work in this direction.

7.1.1 Previous Work and Challenges

The availability of massive labeled datasets and capability of graphic processing unit (GPU) computing has revolutionized the field of deep learning [161]. In 2012, Krizhevsky et al. presented AlexNet, a breakthrough CNN architecture, to classify natural images into 1000 classes, which halved the existing error rate to 16% [66]. Since then, there has been a flurry of successful CNN architectures such as VGG [162], GoogLeNet [163] and ResNet [164] for image classification. ResNets [164] have shown compelling accuracy and convergence behaviors on image recognition tasks. By using identity mappings as the skip connections and after-addition activation, residual units can allow signals to be directly propagated from one block to other blocks.

However, semantic segmentation of images and 3D data is still a nascent area in computer vision, as it is a high level task of understanding and labeling the entire image. Some recent efforts for semantic segmentation with interesting model architectures include fully convolutional nets (FCN) [150], DenseNets [165], U-Nets [166] and Pixelnet [167] for 2D image segmentation; VoxResNet [68] for volumetric brain segmentation and ShapeNet [168] for object classification in non-Euclidean manifolds. V-net [169] uses 3D-CNNs for volumetric medical image segmentation of MRI volumes. They use a series

of convolution layers followed by de-convolution layers, with identity skip connections between convolutional and deconvolutional layers for fine feature forwarding. VoxResNets have demonstrated impressive results on brain segmentation from MRI scans using an architecture with 25 volumetric convolutional layers and 4 deconvolutional layers, making it the deepest 3D convolutional architecture so far [68].

3D segmentation is challenging because of the dataset size causing an exponential memory increase and the intensive computations with 3D convolutions and deconvolutions. We face more challenges due to the dense nature of our data. The 3D datasets for both natural and medical images is sparse, only small regions of the image are of interest, whereas, the microstructures do not contain empty spaces. Current segmentation efforts utilize the state of the art pretrained networks from natural image classification, by extracting features and fine tuning the classification part of the network architecture, which also helps overcome the limited amount of training data [169], [170]. This kind of transfer learning approach [171] is currently not possible for us because we want to use neural networks, trained end-to-end, voxel-to-voxel, that take the microstructure and its physical properties as the input, and predict stress hotspots, a physical phenomenon.

7.2 Dataset

We used the dataset of 56 synthetic microstructures generated in Chapter 5 for HCP materials with Equal CRSS ratio between the three deformation modes. The dataset generation is described in section 5.2.1 and Appendix B. Each microstructure is discretized on a 128x128x128 grid, and each voxel is associated with a set of Euler angles that represent its orientation with respect to the sample Z direction (tensile axis). A uniaxial tensile stress upto a total strain of 2% was applied in Z direction resulting in a spatially resolved stress field.

The stress values in the voxels are averaged grainwise to get rid of simulation artifacts from local instabilities during FFT calculations i.e. Gibbs oscillations. The resultant Von Mises stress distribution is then thresholded above the 90th percentile to designate stress hotspots following the same procedure as in chapters 4 and 5. This results in 10% of the grains designated as stress hotspots. We also experimented with multi-class classification by dividing the grains into 5 classes. The grains were divided into 5 equally spaced stress bins (equal bin sizes). Since this results in some classes having small populations, we also divided the grains into 5 equal frequency bins, each

containing 20 percent of the grain population sorted by stress values. Thus three kinds of voxelwise target data were used: 1) grain-wise stress hotspots, 2) five equally spaced stress bins and 3) five equal population stress bins.

Since the Euler angles represent the grain orientations in a highly nonlinear space with lots of degeneracy, two completely different sets of Euler angles can represent the same orientation. Thus we took the grain-wise features which were found important in chapter 5 and used principal component analysis (PCA) with whitening [172] to reduce dimensions, extracting four principal components which explain 94.5% of the variance. We also tested using fundamental zone quaternions to describe the grain orientation. Thus three kinds of voxelwise input data were used: 1) PCA reduced grainwise features, 2) Euler angles and 3) Quaternions in the fundamental zone.

7.2.0.1 Data Augmentation

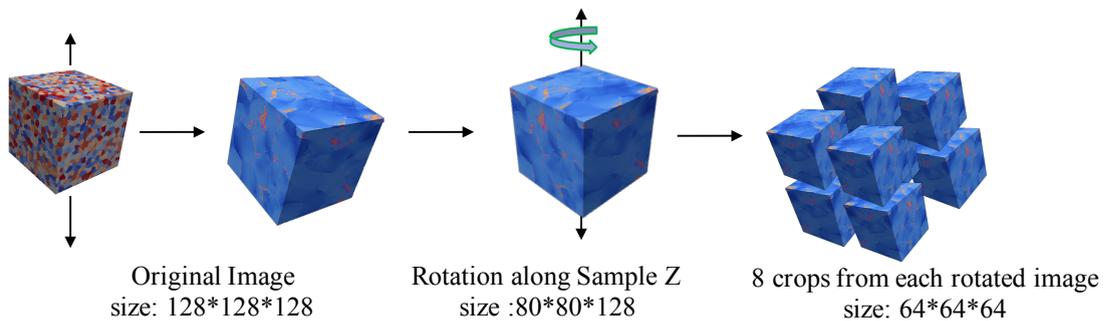


FIGURE 7.2: Data Augmentation Pipeline

Since there are only 56 microstructures, we implemented a data augmentation pipeline as shown in Figure 7.2. It is not physical to rotate the microstructures along the X and Y axes as it changes the grain association with the tensile axis. Hence each microstructure was rotated about the tensile stress axis (Z), and the new grid positions are interpolated. Rotations of r degrees result in $N_{rot} = 360/r$ number of microstructures. Only the central sphere which is free of edge effects is kept in the rotated data set; resulting in microstructures of size $80 \times 80 \times 128$. Each of these microstructures were then sliced into 8 overlapping cubes to get $56 \times N_{rot} \times 8 = 448 \times N_{rot}$ samples. Due to the interdependencies between the Euler axes, linear interpolation in Euler space results in unexpected effects [173]. To avoid this, we used 90 degree rotations about sample Z for data augmentation. For 2D neural networks, to obtain a 2D dataset, image

slices from the 3D microstructure were taken along the Z direction. There are 128 Z slices, and we selected 16 equally spaced slices, as adjacent slices are very similar.

7.3 Model Development and Results

We have attempted both 2D pixelwise segmentation and 3D voxelwise segmentation. With the 3D neural net architectures, we tried out 4 kinds of experiments : voxelwise Hotspot segmentation, voxelwise segmentation into 5-classes based on stress bins, voxelwise regression and grainwise segmentation into 5-classes based on stress bins. With the 2D neural net architectures, we have tried out pixelwise segmentation into 5-classes based on stress bins.

The convolutional layers used the Leaky ReLu [158] activation function and a dropout regularization layer [174] with a rate of 0.8. The output of each convolutional layer was fed to a batch-normalization layer [175], before progressing to the next layer. For regression tasks, the networks try to minimize the Huber loss function [176] and report the model performance with the mean squared error (MSE) metric. For imbalanced binary classification with grainwise stress hotspot segmentation, we use the weighted cross entropy loss function and report the AUC performance metric. For 5-class classification experiments, where we predict Equally spaced stress bins and equal population stress bins, we used the softmax cross entropy loss function and report model performance with mean class accuracy. The model architectures and results are summarized in Table 7.1. We used Tensorflow as the framework for this project, and had access to 3 Tesla K80 GPUs which helped us accelerate training.

We tried out 4 neural net architectures for 3D voxelwise segmentation: Fully convolutional networks (3D-FCN), Convolutional networks with upsampled hypercolumns (3D CNN-hypercolumns), Residual networks with upsampled hypercolumns (3D Resnet) and Residual networks with upsampled hypercolumns and grainwise predictions (3D Resnet-grainwise). For 2D pixelwise segmentation, we have used 2D Residual networks with hypercolumns upsampled from unpooling, with varying depth, and attempted 5-class classification.

7.3.1 3D neural network architectures

We first implemented a 3D fully convolutional architecture (3D-FCN) inspired from Long et al. [150] for voxelwise Hotspot segmentation. After 3 convolutional and 2 max-pooling layers, an input of size $128 \times 128 \times 128$ results in an output of size $32 \times 32 \times 32$, which is upsampled via 2 deconvolutional layers to size $128 \times 128 \times 128$. Then a sigmoid activation is applied to produce class probabilities which are fed to a weighted binary cross entropy loss function. From Table 7.1, we can see that this 3D-FCN network does not converge and has an AUC of 0.51 in validation data. Figure 7.4a shows the predicted output. The output is patterned, which means that the deconvolutional layers are not learning useful representations. High frequency checkerboard-like artifacts are common with using deconvolutional layers which can be overcome with better upsampling strategies [177] such as resizing the image via bilinear interpolation [178] or using sub-pixel convolutions [179]. Since we have a 3D dataset, upsampling using interpolations will be computationally expensive.

Another set of architectures for semantic segmentation use models built on multiscale features extracted from the multiple layers of a CNN, capturing both the low-level and the high-level features. A hypercolumn is constructed by upsampling these multiscale features to input size, and concatenating along the filter dimensions [151], as shown in Figure 7.3b. The hypercolumns are then fed to a multilayer perceptron (MLP) to get the pixelwise predictions. The Pixelnet architecture [167] used hypercolumns computed on-the-fly using sparse upsampling on only 2% of the total pixels in the image, which leads to faster training while avoiding overfitting. It also allows the use of more multiscale features in the hypercolumns, and a multi-layer perceptron (MLP) with more hidden units. During test time, the entire image is upsampled to get dense predictions.

Inspired by this architecture, we implemented a 3D CNN with upsampled hypercolumns (3D CNN-hypercolumns) and a MLP for voxelwise predictions. Instead of using deconvolutional layers or interpolations for upsampling, we implemented a deterministic upsampling approach by undoing the maxpool operations as shown in Figure 7.3a, similar to nearest neighbor interpolation. The hypercolumns are then fed to a multilayer perceptron (MLP) to get pixel-wise classification or regression outputs.

Table 7.1 lists the results of 3D CNN-hypercolumns architecture with PCA reduced features as input and different target tasks. We also tried regression to predict the grain averaged Von Mises stress with voxelwise PCA reduced features as input.



FIGURE 7.3: a) Pooling and unpooling layers. For each pooling layer, the max locations are stored. These locations are then used in unpooling layer [180].¹ b) Hypercolumns from upsampled convolutional layers. Figure reproduced from [151].

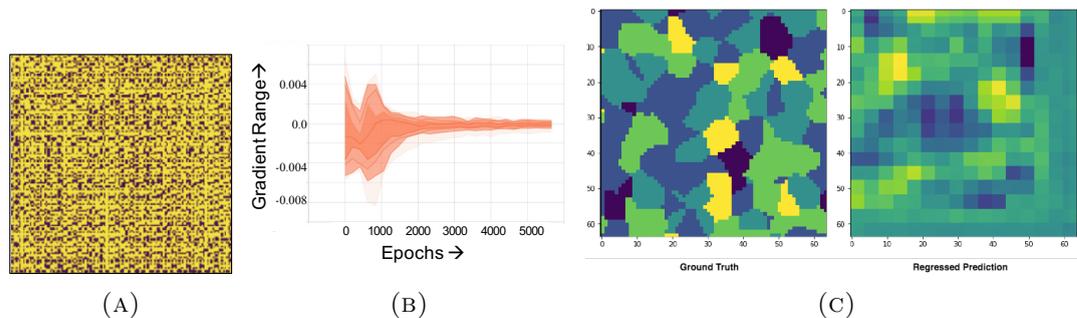


FIGURE 7.4: Problems with 3D architectures. a) Checkerboard like artifacts in 3D-FCN outputs b) Vanishing gradient problems in 3D-CNN-hypercolumns and c) Regression output

Figure 7.4c shows the regression result on validation data after training the network for 100 epochs. We noticed that the regression network tried to minimize the mean squared error over all the voxels, which results in it predicting average stress values. In all cases, the networks stopped training midway due to vanishing gradients as seen in Figure 7.4b. This is a known problem with deep neural networks when using gradient based training schemes during backpropagation [181]. As the number of layers in the neural network increases, training them becomes harder due to vanishing gradients during backpropagation [182], which is overcome through identity skip connections between different convolutional layers that effectively decomposes the deep network into an ensemble of many short networks [164, 183]. Hence we introduced skip connections

¹Reprinted by permission from Springer Customer Service Centre GmbH: Springer Nature, Artificial Neural Networks and Machine Learning – ICANN 2016, “DeepPainter: Painter Classification Using Deep Convolutional Autoencoders”, E. David and N. S. Netanyahu, Copyright 2016.

between different convolutional layers in the 3D CNN-Hypercolumns architecture. The hypercolumns were constructed from the output of all the convolutional layers in the network, which are then fed to a multi-layer perceptron with 1000 hidden units. To utilize the known grain structure of the 3D microstructure, we modified the architecture to predict grainwise probabilities.

This 3D-Resnet-Grain-wise architecture is shown in Figure 7.5. The voxels belonging to each grain in the hypercolumns input are reduced to a single value (maximum or average), to get a hypercolumn feature per grain. This grainwise input is then fed to a multilayer perceptron with 1 hidden layer, to get grain-wise predictions. We used the softmax cross entropy loss function for 5-class classification. During training, sparse upsampling is used for hypercolumn construction, selecting only the indices of the randomly chosen grains for upsampling. Since a 3D microstructure contains ~ 5000 grains, we chose 300 grains ($\sim 6\%$) to upsample. During validation and test times, we upsampled the entire microstructure to get the predictions.

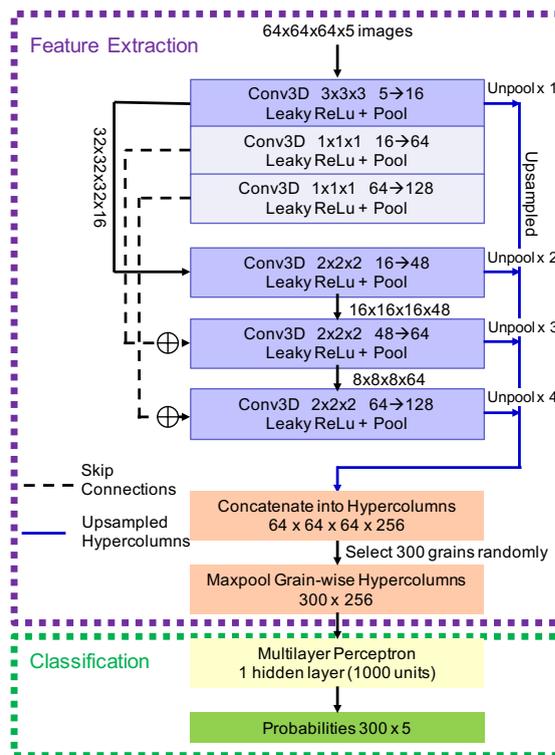


FIGURE 7.5: a) Neural Net architecture for 3D Resnet_Grain-wise model

The input to the network are the voxelwise fundamental zone quaternions and the grain size ($64 \times 64 \times 64 \times 5$). Hence we feed both texture and local geometry information to the network. We predict 5 equal population stress bins, each containing 20 percentile of the grain population. We first trained with network with only 1 original

3D microstructure ($128 \times 128 \times 128 \times 5$), which with data augmentation results in 32 input samples of size $64 \times 64 \times 64 \times 5$. The network was able to overfit the data, hence it has the capacity to learn this dataset. Next, we trained this network with the entire dataset (1224 augmented training samples), using Adagrad optimizer [184], for 50000 epochs. However the network is not able to optimize and the cross entropy loss saturates at a value of 1.6 (Table 7.1). We think this could be due to the nature of our dataset, since we are trying to predict a physical phenomenon from microstructural descriptors, instead of segmenting areas based on pixels. Since two grains, belonging to different classes in a microstructure could have the same quaternion descriptor, it becomes hard for the CNN to distinguish between them. Instead we need to provide spatially correlated data in the input, for the network to function as expected. This can be done by adding voxelwise distance from the grain boundary, triple junctions and quadruple points, as well as the voxelwise kernel average misorientations in the input features. These voxelwise properties can be easily calculated using Dream.3D filters. Another approach would be to separate the microstructure into grains and their neighborhoods, and train the network on these patches to classify the grain. This is a simpler problem than multi-class classification on the whole microstructure image.

7.3.2 2D neural network architectures

A 2D version of the 3D CNN-Hypercolumns with skip connections was implemented as shown in Figure 7.6a. We used 2 versions with 13 and 18 convolutional layers respectively. For constructing hypercolumns, 10000 pixels (24%) were chosen randomly, and the outputs of the convolutional layers was upsampled at those locations. Note that this architecture utilized shorter skip connections as compared to the 3D Resnet with grainwise hypercolumn architecture.

We trained the 2D convolutional neural nets on slices of the 3D microstructure taken perpendicular to the loading direction. We conducted experiments with different kinds of inputs and outputs, which are summarised in Table 7.1. The best performing network used pixelwise Euler angles as the input and classified the pixels into 5 equally spaced stress bins. The network was trained for 170 epochs using Adagrad optimizer [184]. From the confusion matrix (Figure 7.6c), we see that 46% of the pixels belonging to Bin 1 (lowest stress) are predicted as Bin 2 and 3% are predicted as Bin 3. Most of the confusion between classes occurs in adjacent bins. This makes sense, because our

TABLE 7.1: Summary of convolutional neural net development efforts

Architecture	# Conv layers	Input features	Segmentation Target	# Training Samples	Training		Test		Remarks
					Loss	Metric	Loss	Metric	
3-D Architectures									
3D-FCN	3	Euler Angles	Grain-wise stress hotspots	576	0.12	0.51	0.699	0.502	Correlated outputs observed
	3	PCA reduced	Grain-wise stress hotspots	576	0.67	0.51	0.52	0.503	
3D CNN-hypercolumns	6	PCA reduced	Grain-wise stress hotspots	144	0.693	0.518	0.78	0.472	Vanishing gradients and
	6	PCA reduced	5 Equally spaced bins	144	1.05	0.3457	1.34	0.2917	Regression predicts
	6	PCA reduced	Von Mises Stress	144	12.62	251	13.76	243	average stress
3D Resnet	4	Quaternions	5 Equal population bins	24	0.004	99.48%	1.7	16.9%	Can overfit but not
grain-wise	4	Quaternions	5 Equal population bins	1224	1.613	16.17%	1.608	19.35%	generalize
2-D Architectures									
Residual net with hypercolumns	13	Euler Angles	5 Equal population bins	3072	1.06	30.46%	1.17	23.24%	Highest error in Hotspots (Bin4)
	13	Euler Angles	10 Equal population bins	2048	1.61	19.67%	1.62	21.30%	Bad Optimization
	18	Quaternions	5 Equal population bins	128	0.02	98.3%	15.56	17.45%	Can overfit small data
	18	Quaternions	5 Equal population bins	6528	1.61	20.66%	1.60	21.26%	Bad generalization
	13	Euler Angles	5 Equally spaced bins	2048	0.81	60.5%	1.02	54.95%	Best results yet

classes are formed by dividing the grains into bins based on stress thresholds. Hence, the pixels near the boundary of bin threshold have slightly different stress values, and can change bin membership into adjacent bins on adjusting the threshold. We observe that the model never predicts Bin 4 (highest stress). We argue this is because the bins were equally spaced and hence the total pixels in Bin 4 available during training is very small. As we can see, Bin 4 has only 32 out of 4096 pixels in the slice shown.

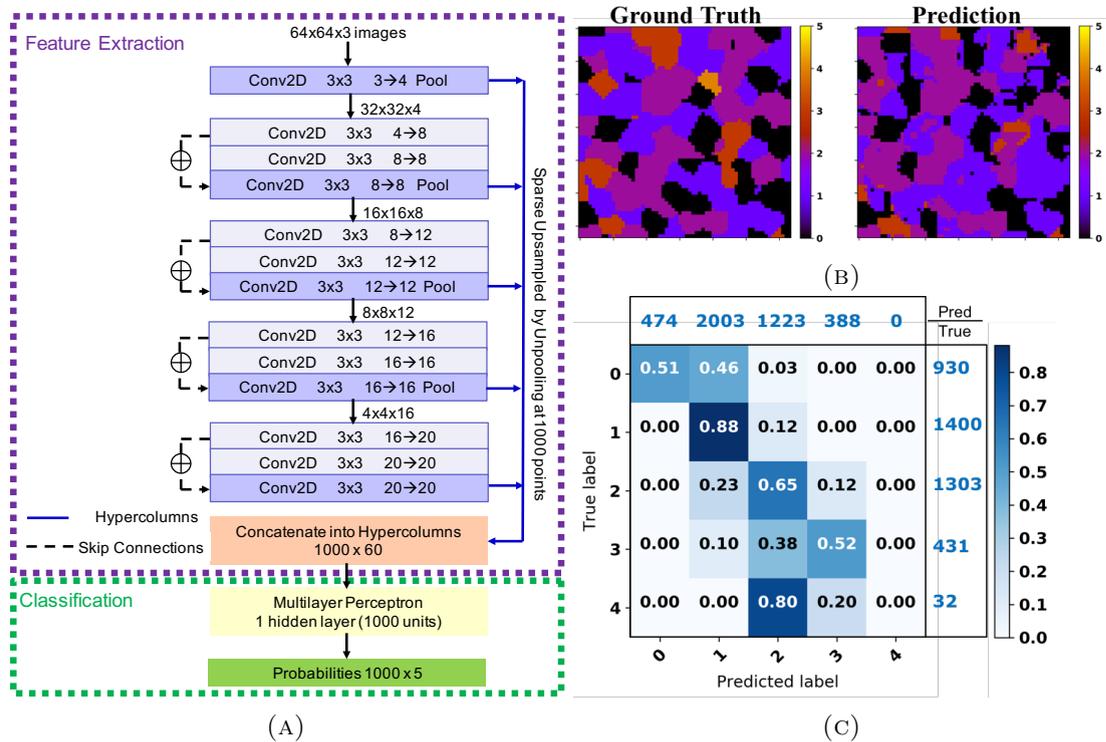


FIGURE 7.6: a) Neural network architecture for a 2D Residual network inspired pixel-wise segmentation model with 13 convolutional layers. b) Ground truth and predicted classes for validation data c) Confusion matrix for validation data. The total number of true labels in each bin are noted at the right, and total predicted labels per bin are noted at the top.

7.4 Summary

We have explored different CNN architectures for voxelwise and pixelwise segmentation of microstructures and have demonstrated the potential for CNNs for microstructure segmentation based on physical phenomenon. During the process we have encountered several roadblocks, which have helped identify the path for applying deep learning methods for microstructure-property relationships. It has been challenging to train CNNs from scratch for this kind of input. CNNs are complex models, and there are many parameters which need to be optimized before successful results. The network has many parameters:

number of convolutional layers, number of filters, filter size, skip connections, number of fully connected units, upsampling method used (interpolations vs. deconvolutions), features used for hypercolumn construction, strategy to derive grain-wise hypercolumns from voxelwise data and training strategy to name a few. The 3D architectures explored here saturate at a cross entropy loss of 1.6. We have achieved moderate success in 2D architectures with voxelwise Euler angles as input and voxelwise 5 equal spaced stress bins as target. Thus, we are positive that tweaking the CNN architectures will help in stress hotspot prediction, and finally help develop multiscale microstructural descriptors.

This exploratory work has helped point out what is missing for applying machine learning techniques to the field of materials science. There are big opportunities for developing standardized methods of representation and working towards a library of microstructural features.

Chapter 8

Future work

In Chapter 4 and 5, data science techniques have been successfully applied to predicting stress hotspots in single phase equiaxed microstructures with different constitutive parameters, for FCC and HCP materials. The natural extension is to predict stress hotspots in more complex microstructures such as dual phase materials (e.g. Ti-6Al-4V alloys) and materials with bimodal grain size distributions. Dream.3D can be used to generate synthetic 2-phase microstructures, and EVPFFT can simulate uniaxial tensile deformation. This dataset can be used to develop more complex microstructural features which contribute to stress hotspot formation. The current dataset can also be evaluated to correlate strain hotspots and slip activities with microstructural descriptors. The approach developed in this work can be extended to experimentally obtained HEDM datasets for understanding physical phenomenon such as annealing, recrystallization and other deformation phenomenon such as twinning. Another direction to pursue would be the use of ensemble learning, where multiple kinds of machine learning model predictions are averaged to get the final prediction.

In Chapter 7, we explored using deep learning techniques. Through the experiments detailed in this chapter, it is clear that there is a need to include spatially correlated information in the input data for the networks to function as expected. The 3D Resnet-grainwise architecture currently samples grains from a single microstructure during each training step, which might not provide enough diversity in the training data. This can be overcome by processing the input to divide each microstructure into patches containing the target grain and its neighborhood, and train the network to classify the central grain. Once a CNN learns the long range feature representations of the grain, then instead of using a MLP on the hypercolumn features, a random forest model can

be used to get an interpretable model with feature insights. Feature selection techniques (from Chapter 3) can be used to tailor the convolutional layers to be selected for the hypercolumn construction. The learnt hypercolumn features can also be studied for interpretability. The important hypercolumn features can be studied to relate them with the input microstructural features by starting with very simple microstructures, and increasing complexity, to see the which hypercolumn features are affected.

The computer vision field has only recently evolved for complete scene understanding from coarse grained to fine grained inference, from image classification to classification + localization and then semantic and instance segmentation. There are different segmentation architectures in use today, with fully convolutional neural nets (FCNs) being the most successful [150]. However, these networks leverage the progress made in the computer vision field for segmentation by transferring the learned features for related tasks by fine-tuning[185] and building the model on top. Similarly, transfer learning has been used to take the features learnt for a simple task (image classification), and apply them to progressively complex tasks [186], [170], [187], [152]. Hence, for voxelwise segmentation to work, it might be necessary to break the segmentation into simpler tasks, modifying the input to do only classification or regression on a related task such as average grain size prediction. Once the network is optimized for the simpler task, it can be progressively optimized for more complex tasks.

Currently the neural networks that use quaternion math are restricted to ones that take quaternion input [188] or predict quaternion output [189]. The recently developed deep quaternion networks [190] provide an interesting research direction for using quaternion weight values. Quaternions can easily represent spatial orientations, and a quaternionic neural network might be able to learn weights representing physical relationships between texture, anisotropy and the stress state. Another interesting avenue to explore would be encoding the microstructure as a transgranular network and using graph based approaches like graph convolutional neural networks [191], [192], [193], [194] and markov random fields [195] which are used for outlier detection tasks.

Data driven techniques are abundant, and there is a strict need for their evaluation and application in the materials science domain, to accelerate the materials development cycle. Collaboration between computer scientists and material scientists can help identify the problems that can be solved using state of the art data driven techniques, and foster endless possibilities of new research directions. The results have demonstrated the use of different data driven techniques for predicting physical phenomena, and have

opened the doors to learning generic microstructural descriptors, which can be applied to learn structure-property relationships in materials. Simulated data can be used to train the models. We must, however, be cognizant of the fact that synthetic data can lead to lower accuracy of developed models. Therefore the goal is not to achieve a particular physical model, but developing an approach for generating microstructural descriptors to understand structure-property relationships.

Appendix A

Constitutive Parameters: Face Centered Cubic materials

The constitutive model parameters for FCC materials represent oxygen free high thermal conductivity (OFHC) copper. The single crystal elastic constants for copper are given in table B.2. FCC materials deform plastically by slip on twelve $\{111\} \langle 110 \rangle$ slip systems. To obtain the actual values of CRSS and the Voce hardening parameters; the Voce model was fit to an experimentally measured stress- strain curve for uniaxial tension in OFHC copper [196] using the VPSC formulation. The results of the fitting are shown in figure B.2. The Voce hardening parameters for this hypothetical case are shown in table B.1. The boundary conditions correspond to uniaxial tension along Z, with an applied strain rate component along the tensile axis $\dot{\epsilon}_{33} = 1s^{-1}$. The EVPFFT simulation was carried out in 200 steps of 0.01%, up to a strain of 4%.

To understand how the most predictive features influence hotspot formation, the distribution of these feature values in normal and hot grains is plotted as shown in figure A.2. It can be seen that the feature distributions for hot and normal grains are different for Schmid factor and the distance of tensile axis (sample Z [001]) from [111], [110] and [100] crystal directions.

TABLE A.1: Single crystal elastic stiffness constants (in GPa) for Copper

Material	C_{11}	C_{12}	C_{44}
Copper	168.4	121.4	75.4

TABLE A.2: Voce Hardening law parameters imitating Copper

CRSS ratio	τ_0^s (MPa)	τ_1^s (MPa)	θ_0^s	θ_1^s
1:1:1	7.43	102.79	356.44	13.01

TABLE A.3: Feature name descriptions

Feature name Abbreviation	Description	Feature name Abbreviation	Description
Schmid_0	Basal $\langle a \rangle$ Schmid factor	100_IPF_x	Distance from the corners of the 100 Inverse pole figure
Schmid_1	Prismatic $\langle a \rangle$ Schmid factor	001_IPF_x	Distance from the corners of the 001 Inverse pole figure
Schmid_2	Pyramidal $\langle a \rangle$ Schmid factor	AvgC_Axes_x	Unit vector components describing the c axis orientation for hcp
Schmid_3	Pyramidal $\langle c + a \rangle$ Schmid factor	Max_mis	Maximum misorientation between a grain and its nearest neighbor
Surface area volume ratio	Ratio between surface area and volume of a grain	Min_mis	Minimum misorientation between a grain and its nearest neighbor
theta	Polar angle of hcp c axis w.r.t sample frame	AvgMisorientations	Average misorientation between a grain and its nearest neighbor
phi	Azimuthal Angle of hcp c axis w.r.t. sample frame	QPEuc	Average distance of a grain to quadruple junctions
TJEuc	Average distance of a grain to triple junctions	NumNeighbors	Number of nearest neighbors of a grain
GBEuc	Average distance of a grain to grain boundaries	Neighborhoods	Number of grains having their centroid within the 1 multiple of equivalent sphere diameters from each grain
KernelAvg	Average misorientation within a grain	FeatureVolumes	Volume of grain
Omega3s	3rd invariant of the second-order moment matrix for the grain, without assuming a shape type	Equivalent Diameters	Equivalent spherical diameter of a grain
Surface Features	1 if grain touches the periodic boundary else 0	AspectRatios	Ratio of axis lengths (ba and ca) for best-fit ellipsoid to grain shape

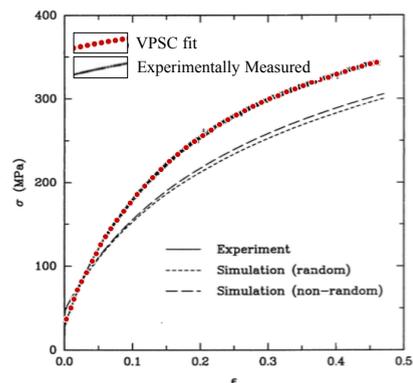


FIGURE A.1: VPS simulation fit to the experimentally observed stress-strain curve for OFHC Copper

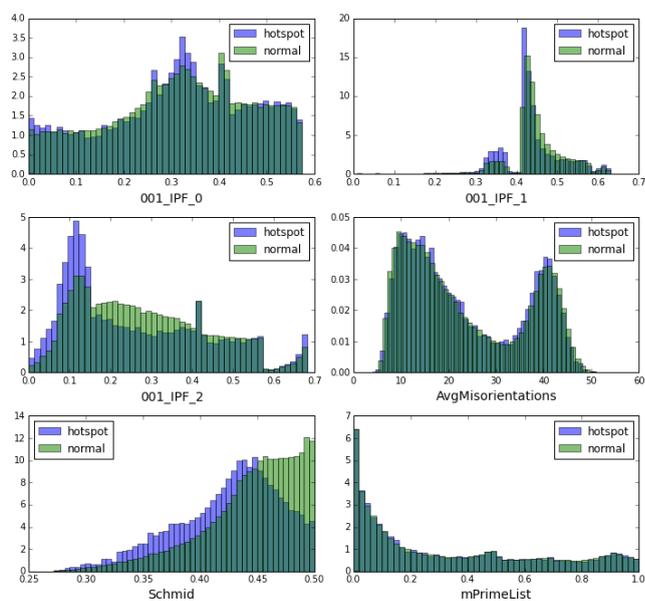


FIGURE A.2: Histogram of the important features distinguishing hot and normal grains in FCC

Appendix B

Constitutive Parameters: Hexagonal Close Packed materials

The constitutive model parameters for HCP materials represent a general alpha-titanium alloy having an equiaxed microstructure. The single crystal elastic constants are given in table B.2. Only three slip systems are considered: basal $\{0001\}[11\bar{2}0]$, prismatic $\{10\bar{1}0\}[11\bar{2}]$ and pyramidal $\langle c+a \rangle$. Two cases are considered based on the strength of different slip systems i.e. having Equal and Unequal CRSS ratios. The Equal CRSS case is hypothetical and is analyzed purely for model development and analysis. The second case with the CRSS ratio of basal $\langle a \rangle$: prismatic $\langle a \rangle$: pyramidal $\langle c+a \rangle = 1: 0.7 : 3$ has the same single crystal elastic stiffness constants (table B.2). The boundary conditions correspond to uniaxial tension along Z, with an applied strain rate component along the tensile axis $\dot{\epsilon}_{33} = 1s^{-1}$. The EVPFFT simulation was carried out in 200 steps of 0.01%, up to a strain of 2%.

To obtain the actual values of CRSS and the Voce hardening parameters; the Voce model was fit to an experimentally measured stress- strain curve for uniaxial tension in α -Titanium [197] using the VPSC formulation. The results of the fitting are shown in figure B.1 and table B.1 lists the CRSS values and hardening parameters obtained for this CRSS ratio. Note that, for HCP materials; we have used 8 different kinds of textures summarized in figure 5.2. The stress exponent is 10 for all cases.

To understand how the most predictive features influence hotspot formation in HCP materials, the distribution of these feature values in normal and hot grains is

TABLE B.1: Voce Hardening law parameters for α -Titanium

CRSS ratio	Slip System	τ_0^s (MPa)	τ_1^s (MPa)	θ_0^s	θ_1^s
0.7:1:3	Basal	82.8	36.7	406.3	4.6
	Prismatic	57.9	25.7		
	Pyramidal	248.5	110.1		
1:1:1	All	100	50	500	10

TABLE B.2: Single crystal elastic stiffness constants (in GPa) for α -Titanium

Material	C_{11}	C_{12}	C_{13}	C_{33}	C_{44}	C_{66}
α -Titanium (approx.)	170	98	86	204	51	66

plotted as shown in figure B.3a for Equal CRSS and figure B.3b for Unequal CRSS ratio.

Equal CRSS: From the plot for ‘theta’ for hot grains, we can see that there is a peak at high ‘theta’ values, where the elastic modulus is low i.e. undergoing more plastic deformation. There are three more smaller peaks at near 64, 40 and 8 degrees, which might be due to more complex effects of plasticity.

Unequal CRSS: the distribution of these feature values in normal and hot grains is plotted as shown in figure B.3b

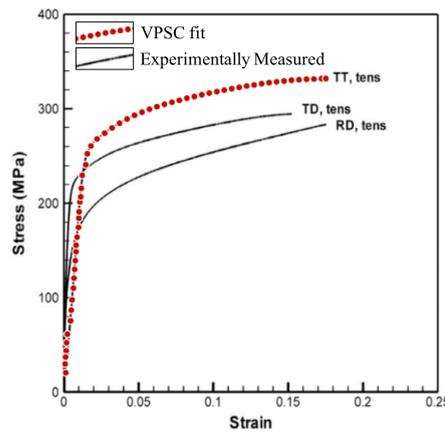
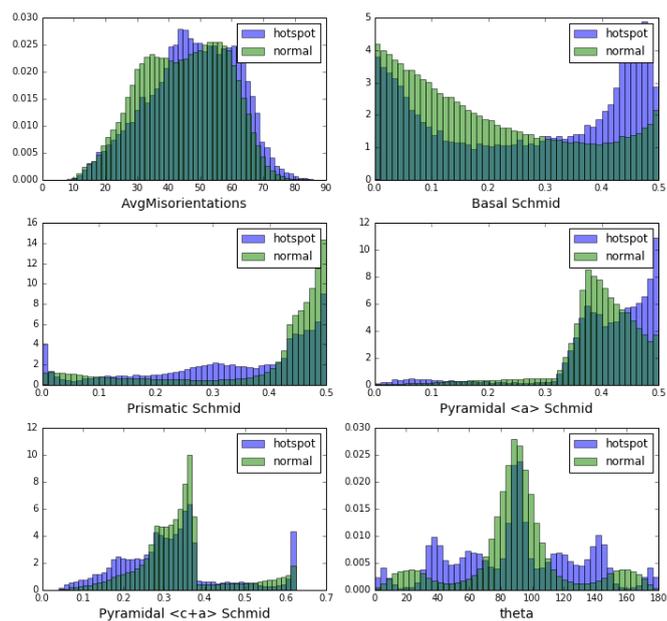
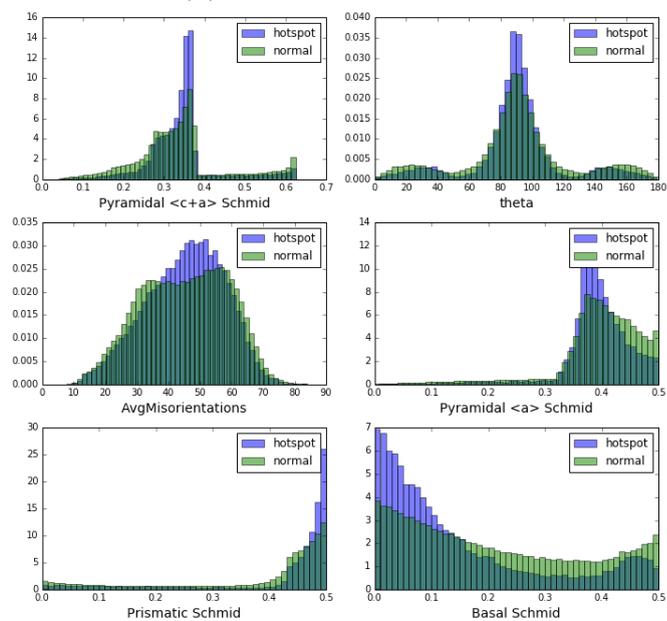
FIGURE B.1: HCP α -Titanium fit

FIGURE B.2: VPSC simulation fit to the experimentally observed stress-strain curve for alpha-Titanium.



(A) HCP Equal CRSS case



(B) HCP unequal slip case

FIGURE B.3: Histogram of the important features distinguishing hot and normal grains

Appendix C

Understanding Hotspot Neighborhoods

Single point statistics such as the average grain size in a polycrystal are effective indicators of the microstructure properties, but in polycrystals, long range interactions between individual grains and grain boundaries impact the microstructural properties. Local neighborhood and boundary mobility distribution has been shown to affect the frequency of abnormal grain growth [142]. The grain boundary network topology is important for development of high temperature superconductors, where current passes when a network of grains with low angle misorientations exists [198]. Grain boundary engineered materials have high fraction of special boundaries such that they resist intergranular crack propagation [88, 143], intergranular corrosion [144] and liquid metal embrittlement[145]. The network topology affects some properties more directly than others [199], and the effect of network topology on stress concentrations has not been studied yet. Hence we propose using microstructural descriptors based on grain connectivity to understand the effect of long range interactions and grain topology on stress hotspot formation.

C.1 Transgranular network

Both the grain boundary and the transgranular network have the same topological and crystallographic information, but the representation chosen for analysis is based on the problem at hand. For example, grain boundary network is used to study intergranular corrosion cracking [200], where damage propagation occurs along a contiguous

grain boundary path. Transgranular network has been used to study super-conducting polycrystals[201] and abnormal grain growth [72] because these problems involve transport across grain boundaries.

Grain averaged representation of stress hotspots naturally leads to a transgranular network representation of a microstructure. This is an undirected graph G , whose vertices (V) correspond to each grain in the polycrystalline system, and the edges ($E \subseteq V \times V$) connecting the contiguous grains (nodes) represent the grain boundaries. This is the dual of the grain boundary network, as shown in figure C.1. In *labeled* graphs, the vertices (nodes) and edges can be associated with discrete or continuous attributes, so called *labels*. A *walk* is a sequence of vertices $w = v_1, v_2, v_3 \dots v_{n+1}; v_i \subseteq V$ traversed in the graph such that an edge $(v_i, v_{i+1}) \in E$. The length of the walk is the number of edges in this sequence. A *path* is a walk in which $v_i \neq v_j \Leftrightarrow i \neq j$ i.e. no vertex appears twice. A *cycle* is a path with $(v_{n+1}, v_1) \in E$ i.e. a closed circuit walk where the walk starts and ends at the same node. A graph $G' = (V', E')$ is a *subgraph* of another graph $G = (V, E)$ iff $V' \subseteq V$ and $E' \subseteq E \wedge ((v_1, v_2) \in E' \rightarrow v_1, v_2 \in V')$. A *clique* is a subset of nodes that are all connected to each other i.e. form a fully connected graph. Cliques of size 3 (i.e. a triangle) form triple lines in 3-D microstructures and triple points in 2-D microstructures. Cliques of size 4 only exist in 3-D microstructures and correspond to quadruple points.

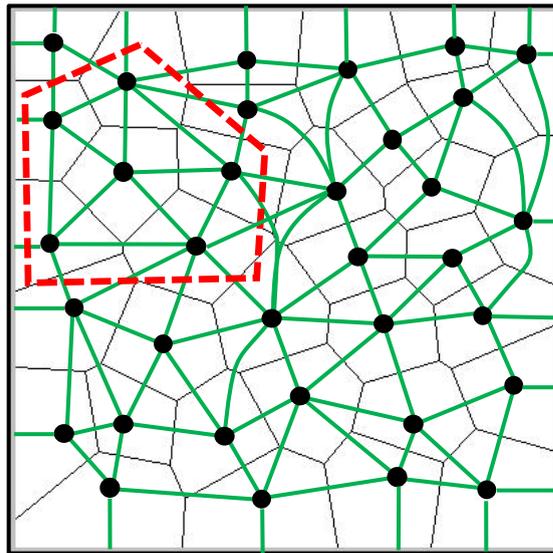


FIGURE C.1: A 2-d microstructure. The grain boundary network is shown in black. It's dual, the transgranular network (shown in green) is constructed with grains as the nodes (drawn at grain centroids), and connections across each grain boundary. An ego graph of radius 1 (only first nearest neighbors) is highlighted in red.

We choose the transgranular representation for analyzing microstructure neighborhood effects causing stress hotspots as it can directly model the crystallographic relationships between a grain and its surrounding neighbors. The nodes in the transgranular network can be labeled with grain properties such as grain orientations, crystal structures, phases or the alignment of tensile axis with easy/hard directions i.e. amenability to deformation. The edges can be labeled with grain boundary properties such as diffusion coefficients, crystallographic misorientation, boundary curvature, mobility and energy by the misorientation between grains.

The local microstructure neighborhood can be described by a subgraph derived from the transgranular network. These subgraphs as known as *ego graphs*; an ego graph of radius n consists of the grain in question, and its neighboring grains until n^{th} order. Figure C.1 shows an ego graph of radius 1. The impact on stress hotspots is studied by comparing the subgraphs between grains in the following sections.

C.2 Comparing Grain Neighborhoods

Following the work of Akoglu et. al [146], the ego graph for each grain in the transgranular network is computed. For each egonet, the number of edges, nodes, sum of edge weights and principal eigenvector and eigenvalues of the adjacency matrix are computed. These form grain-level features, and can be used alongside crystallography based features for stress hotspot prediction.

From the dataset of Unequal CRSS HCP materials developed in chapter 5, we choose a microstructure with the representative initial structure shown in Figure C.2. We construct transgranular networks on this microstructure, with unlabeled nodes and edges labeled by the misorientation between the grains.

Akoglu et. al [146] developed rules for neighborhoods of real graphs, to help pinpoint the nodes with anomalous neighborhoods. The **density power law** [146] states the power law between the number of nodes (N_i) and the number of edges (E_i) of the neighborhood graphs:

$$E_i \propto N_i^\alpha, 1 \leq \alpha \leq 2 \quad (\text{C.1})$$

where i is the i^{th} egonet G_i for a node i of graph G . Figure C.3 shows the density power law plots for neighborhoods of hot and normal grains. The slope, α for

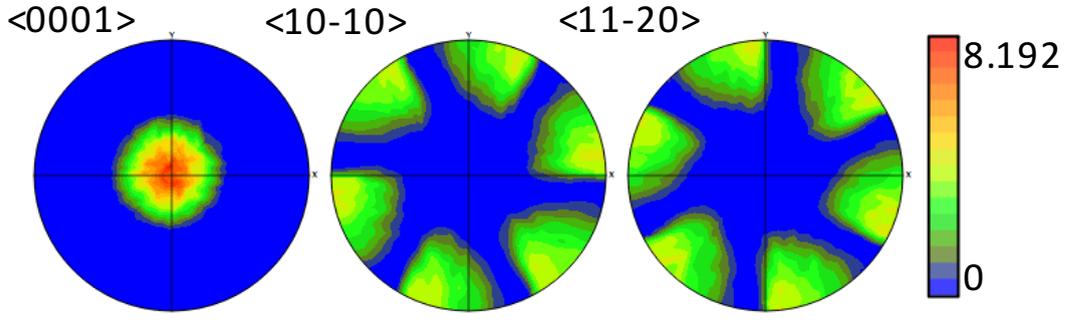


FIGURE C.2: Pole figure representing the texture of the microstructure chosen for analysis using spectral features. The tensile direction is Z (out of plane)

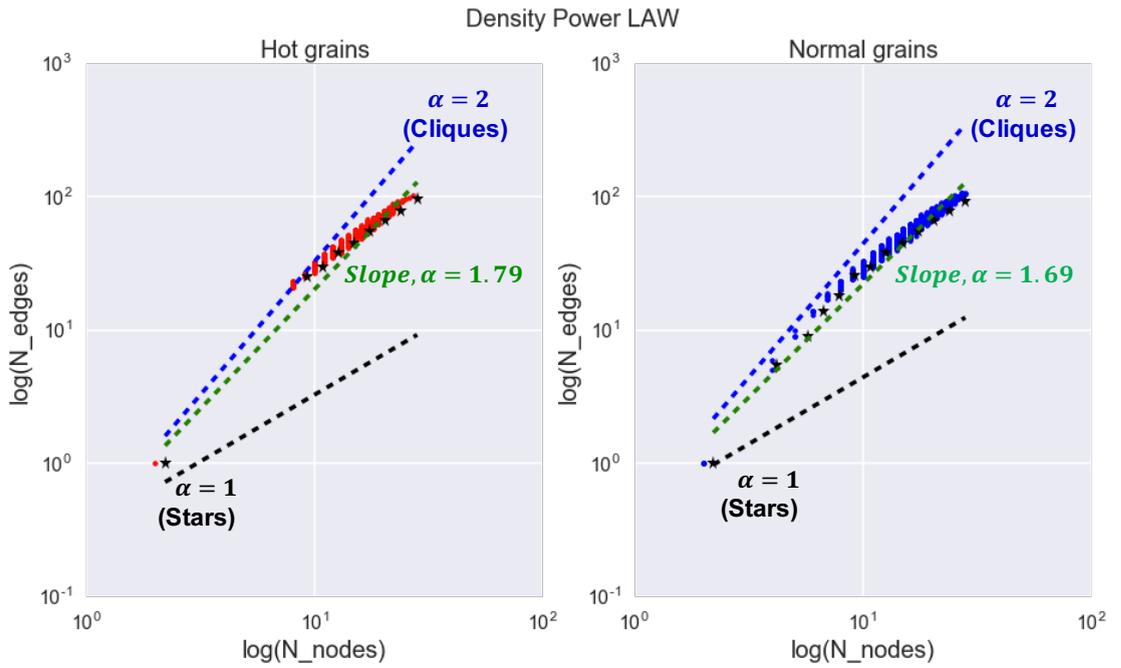


FIGURE C.3: Density power law. Plots show total count of nodes vs. total count of edges in the egonet for all nodes (in log-log scales). The green line is least squares fit on the median values (black stars). Hot grains have more clique like neighborhoods.

the neighborhood graphs of hot grains is higher than normal grains, which means that the highly stressed grains have more number of edges i.e. are more connected.

According to the **weight power law**, the sum of edge weights (W_i) and the number of edges (E_i) of the neighborhood graphs follow a power law:

$$W_i \propto E_i^\beta, \beta \geq 1 \quad (\text{C.2})$$

Figure C.4 shows the weight power law plots for neighborhoods of hot and normal grains. The slope, β for the neighborhood graphs of hot grains deviates from the weight power

law equation C.2.

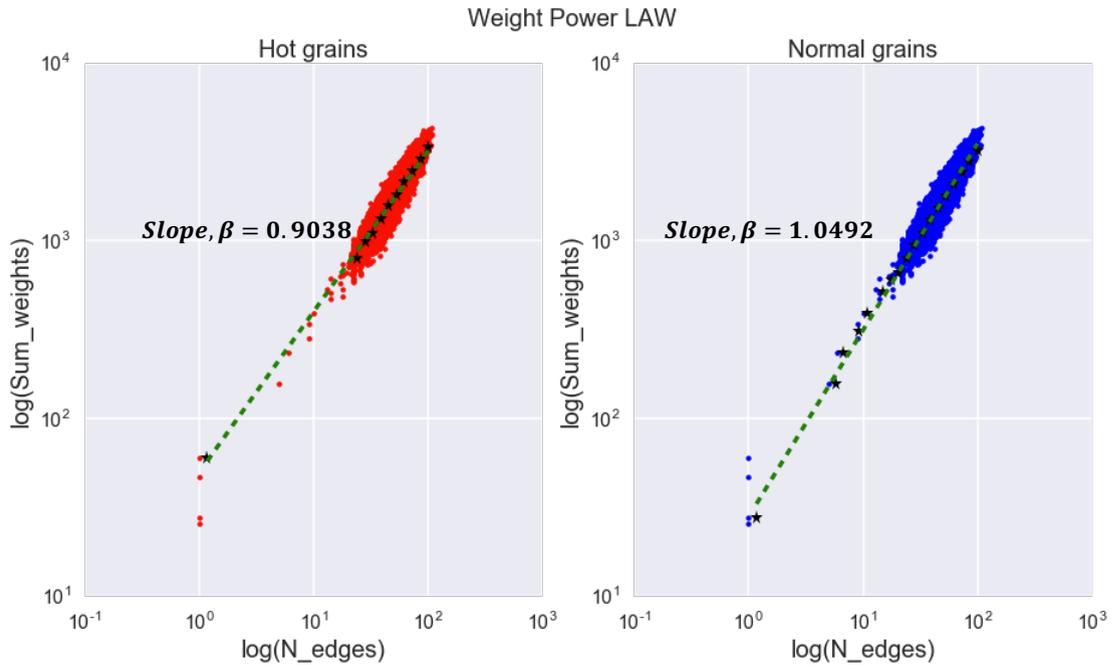


FIGURE C.4: Weight power law. Plots show total weight vs. total count of edges in the egonet for all nodes (in log-log scales). The green line is least squares fit on the median values (black stars). The neighborhood graphs for hot grains deviate from equation C.2

The **eigenvector power law** [146] states the power law between the principal eigenvalue of weighted adjacency matrix (λ_i) and the total weight of an ego-graph (W_i):

$$\lambda_{w,i} \propto W_i^\gamma, 0.5 \leq \gamma \leq 1 \quad (\text{C.3})$$

Figure C.5 shows the eigenvector power law plots for neighborhoods of hot and normal grains. The slope, γ for the neighborhood graphs for hot grains is outside the range specified by equation C.3. The normal grains have a uniform weight distribution as γ is close to 0.5.

C.2.1 Effect of grain neighborhood on stress hotspots

Stress hotspots by themselves do not form a connected network (D), neither are they spatially correlated as seen in section 4.2.1.4. However, the local neighborhood of hot grains is different from normal grains as witnessed by the different power law constants in the previous section. To quantify the influence of the neighbors, the grain neighborhood ego graphs of radius 3 are studied. An ego graph of radius n is a subgraph of the network built on a node and its nearest neighbors from 1st to the nth order.

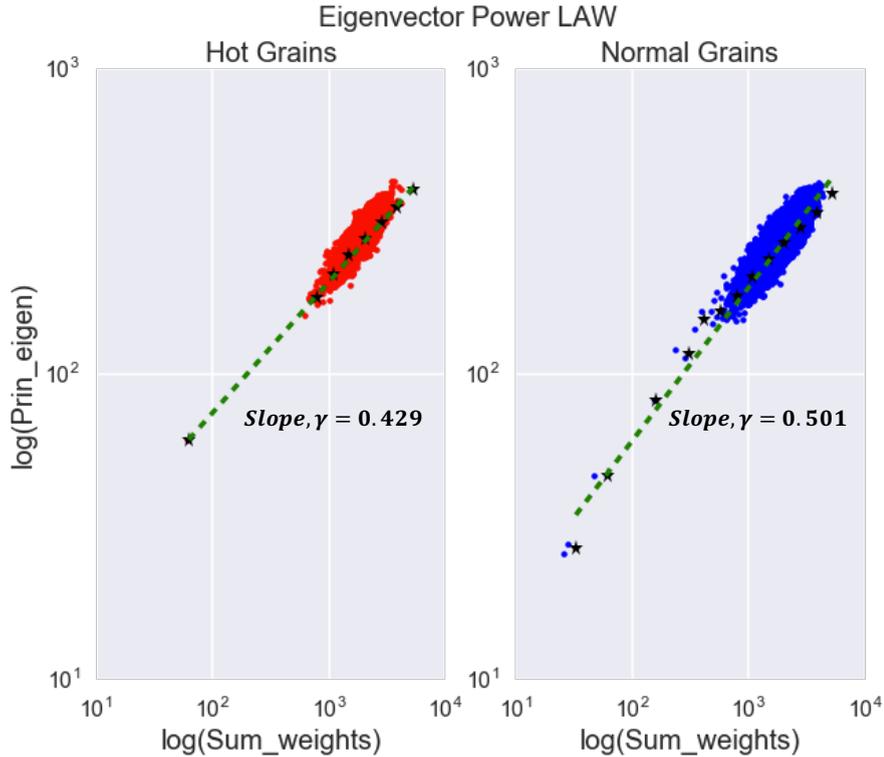


FIGURE C.5: Eigenvector power law. Plots show the principal eigenvalue vs. total sum of weights in the egonet for all nodes (in log-log scales). The green line is least squares fit on the median values (black stars).

The relationship between the probability of a grain and its n th nearest neighbor being a hotspot is studied. Using the random forest models developed for Unequal CRSS ratio HCP materials, we can predict the probability of each grain being a hotspot. To understand the effect of neighbors, we construct a list of the probability of hotspot formations ($P_{hotspot}$) for the first, second and third nearest neighbors of each grain. On an average, in equiaxed polycrystalline microstructures consisting of ~ 5200 grains; each grain has 14.67 first nearest neighbors, 222.93 second nearest neighbors and 3377.10 third nearest neighbors. For each class of neighbors, a list of the probability of hotspot formation ($P_{hotspot}$) is calculated. Then for each grain, the following metrics are calculated from the list for each class of neighbors to describe the hotspot potential: the mean, standard deviation, sum, maximum probability, minimum probability, and the count of the number of n^{th} neighbors whose probability of being a hotspot is > 0.1 , since 10% of the grains have been designated as hotspots and the random chance of a hotspot is 0.1. Table C.1 lists the mean value of these features over 80 different microstructures. Figure C.6 shows the distribution of these features between the two classes. It is observed that the features for the second nearest neighbors have a slightly different distribution.

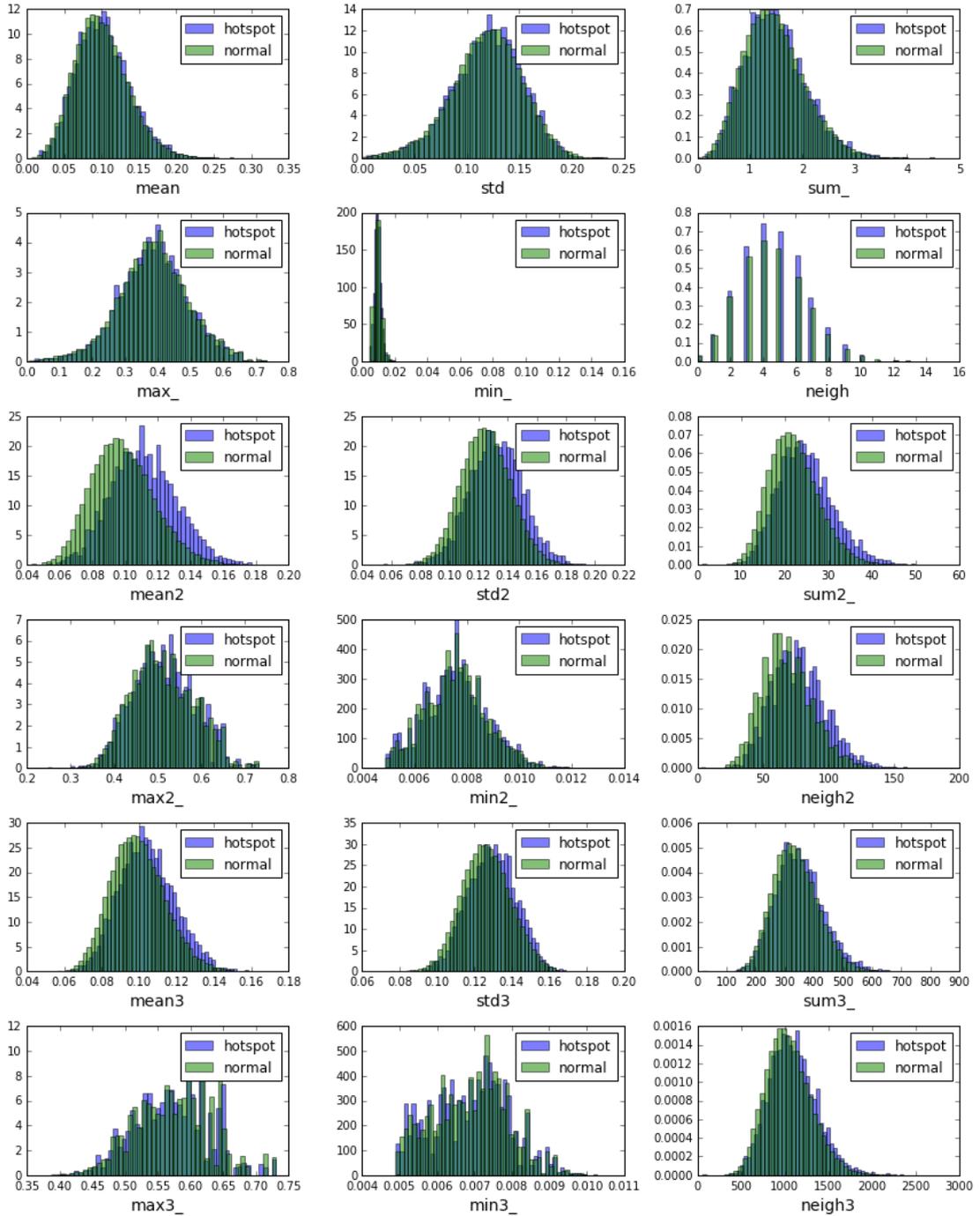


FIGURE C.6: Histogram of the neighbor probability based features distinguishing hot and normal grains for 80 different kinds of microstructures

TABLE C.1: Mean probability based feature values for n^{th} nearest neighbors

Neighbor Order	Feature	Mean Value
1 st	mean_1	0.099
	std_1	0.118
	sum_1	1.467
	max_1	0.386
	min_1	0.009
	neigh_1	4.543
2 nd	mean_2	0.099
	std_2	0.125
	sum_2	22.288
	max_2	0.512
	min_2	0.008
	neigh_2	69.037
3 rd	mean_3	0.099
	std_3	0.127
	sum_3	337.71
	max_3	0.573
	min_3	0.007
	neigh_3	1045.95

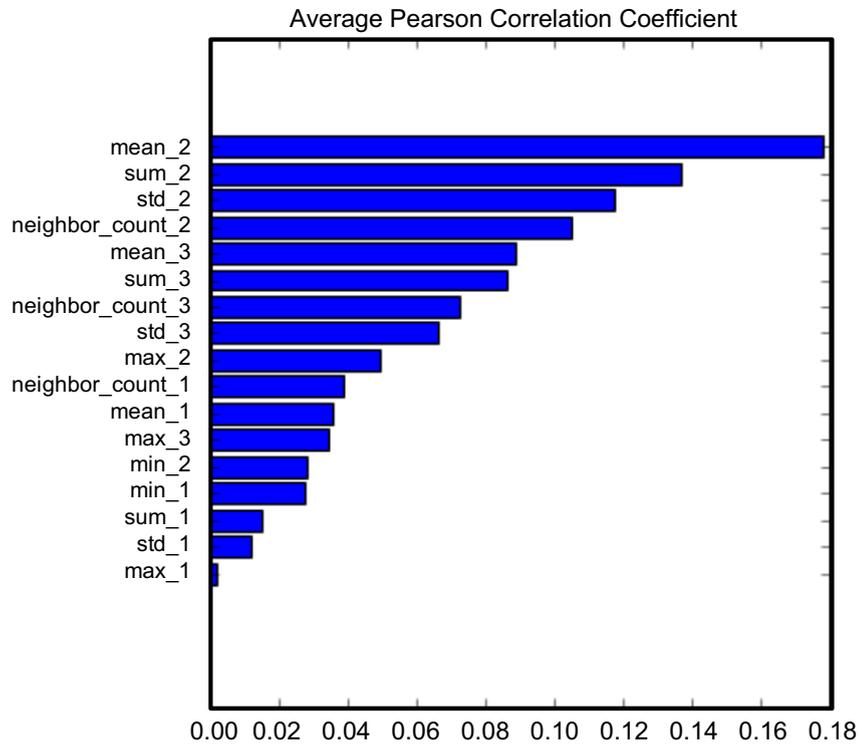


FIGURE C.7: Average absolute Pearson correlation coefficient averaged over 80 different microstructures. The features derived from second nearest neighbors stack together at the top.

The Pearson correlation coefficient is calculated for these features being a hot grain. Figure C.7 shows the average absolute Pearson correlation coefficient for significant features averaged over 80 different microstructures. The features derived from second nearest neighbors stack together at the top and hence play an important role in hotspot determination. The correlation between hot grains and their first nearest neighbors is minimal. The correlation is positive for all except that it varies between positive or negative for standard deviation features. Hence the second nearest neighbors of a hot grain are more probable to be hot themselves.

C.3 Summary

We have compared hotspot neighborhoods using ego graphs derived from the transgranular network. In section C.2, the power laws for the neighborhood graphs have demonstrated that the neighborhood of hot grains is different from the normal grains.

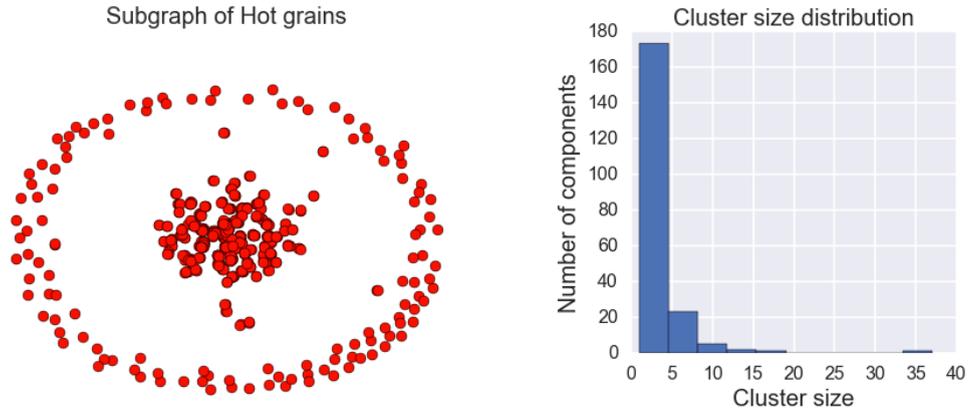
The relationship between the probability of a grain and its n^{th} nearest neighbor being a hotspot is studied in section C.2.1. It was found that the second nearest neighbors of a hot grain are more probable to be hot themselves (figure C.6 and C.7). This finding hints at some kind of medium range order between stress hotspots [148] and further analysis is required to arrive at robust conclusions.

Appendix D

Characterizing the Transgranular Network

The theory of self-organized criticality (SOC) [202] for real life graphs argues that networks representing real physical phenomena exist just near their ‘critical point’. A rapid change in behavior occurs near this critical point. For example, in percolation theory, there exists a critical percolation threshold p_c , when a fully connected path occurs in a system, thus making it ‘percolating’ [203]. This kind of a critical parameter also exists in other naturally occurring systems. For example, fractals are characterized by a motif, and the pattern is similar to scaled copies of that motif; in freezing water, the crystal size distribution is characterized by a power law. Phase changes in materials also occur at a critical point, for example freezing and boiling points represent this phenomenon. Thus, real world networks must exist just above this critical point in a ‘just connected’ state. For example, in a communication network, information flow is achieved at the critical point, and adding any more connections would be suboptimal.

In section C.1, the transgranular network representation of the microstructure was discussed. The robustness of a network is its ability to withstand internal and external failures and perturbations. This is a critical attribute for many systems represented by networks, such as social networks, recommendation systems, computer communication networks, gene/protein biological networks [204] and colloidal gels[205]. The network robustness is characterized by the number of connections between nodes. If there are enough connections between the nodes, the network structure will be robust, and might represent the load bearing part of the microstructure such as in colloidal gels [205]. This idea is studied using percolation theory.



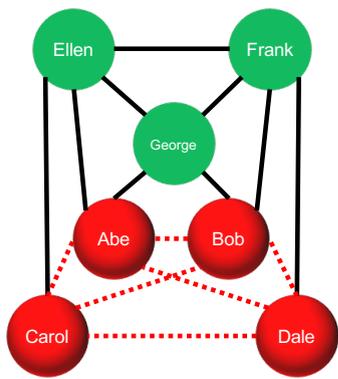
(A) Subgraph induced on hot grains in a transgranular network (B) Cluster size distribution in the subgraph.

FIGURE D.1: Cliques containing stress hotspots in a transgranular network form a percolating network.

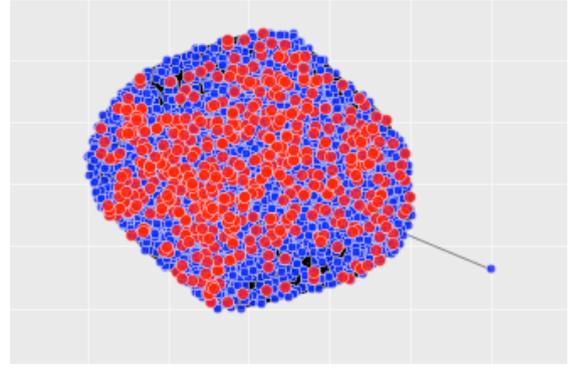
Finding a load bearing network in a material amounts to finding a rigid network, which is related to the rigidity percolation problem [206, 207], different from the more common concept of connectivity percolation. Finding a rigid network is complex, as it depends on the structural details of the node connectivity, and not only on just finding a connected path. The rigidity percolation problem in 2-dimensions has been studied extensively for network glasses [208, 209], by counting the number of floppy modes. The concepts from graph rigidity have been successfully applied to detect the rigidity threshold in 2-D networks [210]. A fully connected network is rigid and over-constrained. As the nodes and edges are deleted, the network becomes diluted, and the transition from a floppy to a rigid network defines the rigidity percolation threshold. The rigid network consists of a ‘spanning over-constrained region’ which forms the stress carrying backbone [211]. However, characterization of load bearing networks in three dimensions remains an open problem.

For transgranular networks, the first attempt at trying to identify a rigid network is by looking at the subgraph formed by stress hotspots. The hot grains, being only 10% of the nodes, do not form a connected network. Figure D.1a shows the subgraph induced on the hot grains, and figure D.1b shows the distribution of cluster sizes in this subgraph. The majority of the subgraph consists of isolated nodes, and the size of the largest connected component is 35 nodes. Hence a different approach is used to find a connected network.

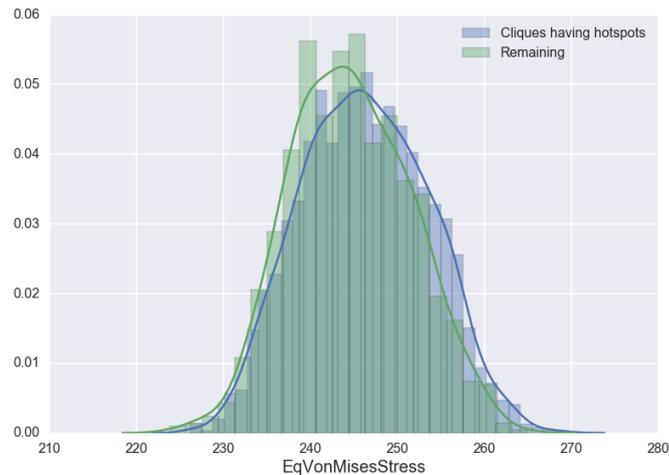
Cliques are maximal connected subgraphs i.e. a group of nodes that are directly connected to all the nodes in that group. Figure D.2a highlights a clique in a toy social



(A) Clique in a social network



(B) Transgranular network where the grains belonging to cliques containing the hot grains are colored red



(C) Stress distribution in the subgraph formed by cliques containing hotspots and remaining grains

FIGURE D.2: Cliques containing stress hotspots in a transgranular network form a single connected component.

network. The cliques containing hot grains are calculated and shown in figure D.2b. This subgraph consists of about 76% of the total number of grains in the microstructure and is connected i.e. forms one giant cluster. Figure D.2c shows the stress distribution in the grains belonging to the subgraph and the remaining microstructure. The average stress in the subgraph grains is higher than the remaining grains. These grains could be related to the load bearing network in the microstructure.

In the context of transgranular networks, two different critical points need to be investigated: the connectivity (percolation), and the rigidity threshold. In 2-D networks, the rigidity threshold lies above the connectivity threshold, and the rigid network consists of the stress carrying backbone. In 2-D, the rigidity threshold has been characterized by comparing the network subgraphs to Laman graphs which describe the minimally rigid systems of rods and joints in the plane [212]. However, the concept has not been

TABLE D.1: Connectivity percolation thresholds for site and bond percolation in different lattices and dimensions [2]

Lattice	Number of nearest neighbors	Site percolation	Bond percolations
1d	2	1	1
2d Honeycomb	3	0.6962	0.65271
2d Square	4	0.592746	0.5
2d Triangular	6	0.5	0.34729
3d Diamond	4	0.43	0.388
3d Simple cubic	6	0.3116	0.2488
3d BCC	8	0.246	0.1803
3d FCC	12	0.198	0.119
4d Hypercubic	8	0.197	0.1601
5d Hypercubic	10	0.141	0.1182
6d Hypercubic	12	0.107	0.0942
7d Hypercubic	14	0.089	0.0787

generalized to three dimensional structures and there exist no metrics in the literature to quantify the rigidity of the 3d network as the underlying mathematical theory becomes invalid [213, 214].

Table D.1 lists the connectivity percolation threshold for site and bond percolation for various lattices in different dimensions. It can be observed that within a given dimension, the percolation threshold decreases with increasing number of nearest-neighbors. Since the 3-D grain structures studied in this work have an average of 14.67 nearest neighbors, the connectivity site percolation threshold should be less than 0.198.

Based on this notion, we start ‘shattering’ the graph [215] by using ‘Islands in the network’ approach. The Islands in the network function [216] takes a graph, and applies a threshold (“water level”), letting all edges or nodes above a certain value through, and removing all others. At a certain threshold, when all links/nodes with that value are dropped, the transgranular network disintegrates into a number of smaller, disconnected clusters. Above this critical fraction, there exists a giant connected component, spanning through the entire microstructure. This critical fraction corresponds to some sort of a phase transition or a percolation threshold.

Since we are trying to find the load-bearing network, it must consist of grains with high stress. Hence, the nodes in the transgranular network are thresholded on stress values, and the remaining nodes are used to find the connected components (the clusters). At each stress threshold, the nodes below the threshold stress are deleted

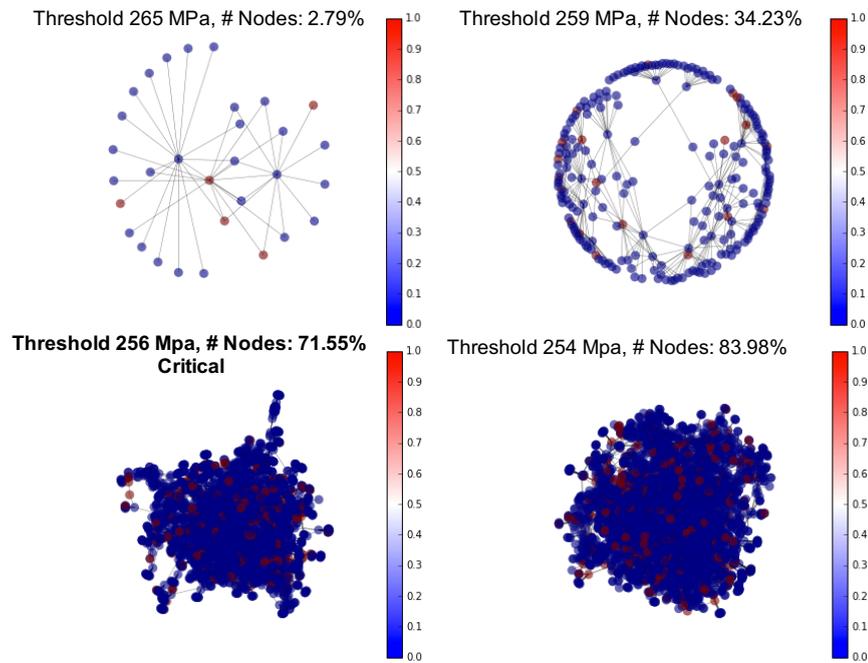


FIGURE D.3: The largest connected component in the transgranular network for different stress threshold values. The red nodes are stress hotspots

and the graph structure is studied. The resultant largest connected component in this thresholded graph for some stress thresholds are shown in figure D.3. The hot nodes in the transgranular network are colored red. The stress threshold and the percentage of total nodes in the giant connected component are indicated. The critical stress threshold is 256 MPa and the corresponding giant connected component has 71.55% of the total grains in the microstructure. Above this threshold, a single giant component spans the microstructure. Below this threshold, the graph consists of multiple clusters, each with its own mechanical properties. Since this threshold is way above the connectivity threshold for site percolation in 3D, this might be related to the rigidity transition.

To find out the critical stress threshold, the largest cluster size and the number of clusters is plotted against the stress threshold in figure D.4. It can be observed that the network undergoes a phase transition, from a set of small, disconnected clusters to a giant cluster (connected component), consisting most of the nodes of the network. Appel et. al. [215] have shown that the critical point is best determined from the effective diameter vs. number of edges in graph. From figure D.4, we can see that the critical diameter for this graph is 72 at 3057 edges. The critical threshold is shown by a vertical line. In figure D.4 only one critical point is observed, which is well above the connectivity percolation threshold. It is hypothesized that this point might correspond to the rigidity

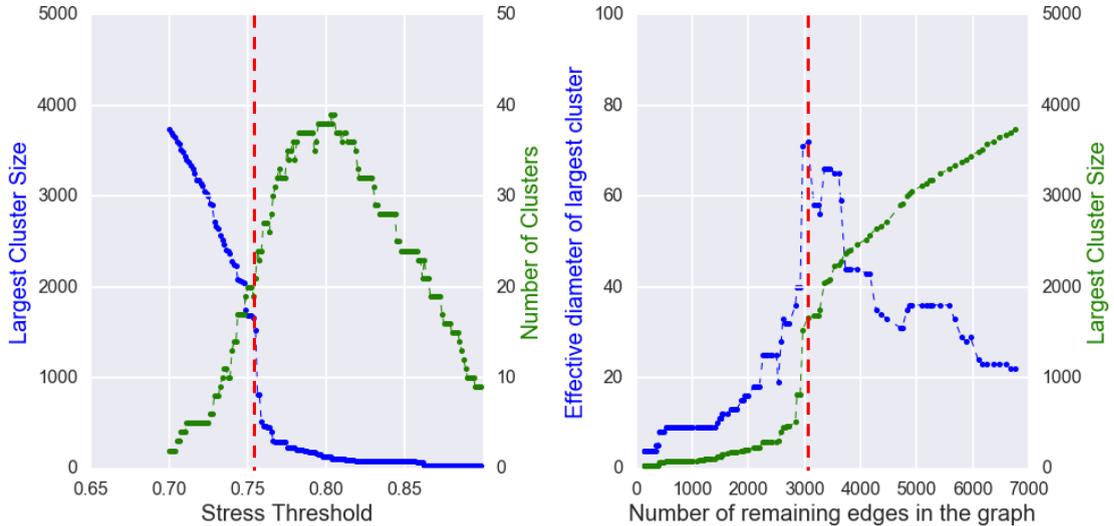


FIGURE D.4: Nodes below the threshold stress are deleted and the graph structure was studied. The critical point for connectivity percolation is shown by a vertical line.

transition because the giant connected component consists of highly stressed grains, an indicator of the ‘stress carrying backbone’ in the microstructure.

D.1 Summary

In this chapter, the transgranular representation of the microstructure has been used to study the relationship of stress hot spots with the load bearing network, and understand the effect of grain neighborhood on hotspot formation.

It has been shown that the rigid load bearing network in a material should consist of a spanning over-constrained region which forms the stress carrying backbone [211]. The subgraph induced on the stress hotspots is not connected and hence disregarded as a load bearing network. The subgraph formed by the cliques containing the hot grains forms one giant connected component and consists of about 76% of the total number of grains in the microstructure. It is yet to prove if this forms a rigid network.

The grains are thresholded on stress values and the induced subgraphs are studied. A phase transition like event happens at 71.55% remaining node as seen in figure D.3 and D.4. It is hypothesized that this point might correspond to the rigidity transition because the giant connected component consists of highly stressed grains, an indicator of the ‘stress carrying backbone’ in the microstructure. However, further analysis needs to be conducted to prove this hypothesis.

Bibliography

- [1] David Mercier. Strain Transfer Across Grain Boundaries — Slip transfer analysis toolbox 2.0.0 documentation, 2013. (Cited on pages [x](#) and [37](#).)
- [2] Kim Christensen. Percolation theory, 2002. (Cited on pages [xiii](#) and [110](#).)
- [3] National Science and Technology Council (US). *Materials genome initiative for global competitiveness*. Executive Office of the President, National Science and Technology Council, 2011. (Cited on page [1](#).)
- [4] Jordan O’Mara, Bryce Meredig, and Kyle Michel. Materials Data Infrastructure: A Case Study of the Citrination Platform to Examine Data Import, Storage, and Access. *Jom*, 68(8):2031–2034, 2016. (Cited on pages [1](#) and [63](#).)
- [5] Alden Dima, Sunil Bhaskarla, Chandler Becker, Mary Brady, Carelyn Campbell, Philippe Dessauw, Robert Hanisch, Ursula Kattner, Kenneth Kroenlein, Marcus Newrock, Adele Peskin, Raymond Plante, Sheng Yen Li, Pierre François Rigodiat, Guillaume Sousa Amaral, Zachary Trautt, Xavier Schmitt, James Warren, and Sharief Youssef. Informatics Infrastructure for the Materials Genome Initiative. *Jom*, 68(8):2053–2064, 2016. (Cited on pages [1](#) and [63](#).)
- [6] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.*, 1(1), 2013. (Cited on page [1](#).)
- [7] Ghanshyam Pilania, Chenchen Wang, Xun Jiang, Sanguthevar Rajasekaran, and Ramamurthy Ramprasad. Accelerating materials property predictions using machine learning. *Sci. Rep.*, 3:2810, 2013. (Cited on page [1](#).)
- [8] Richard J. Needs and Chris J. Pickard. Perspective: Role of structure prediction in materials discovery and design. *APL Mater.*, 4(5), 2016. (Cited on page [1](#).)

- [9] Krishna Rajan. Materials Informatics: The Materials " Gene " and Big Data. *Annu. Rev. Mater. Res.*, 45:153–69, 2015. (Cited on pages 1 and 30.)
- [10] Julia Ling, Max Hutchinson, Erin Antono, Sean Paradiso, and Bryce Meredig. High-Dimensional Materials and Process Optimization using Data-driven Experimental Design with Well-Calibrated Uncertainty Estimates. *Integr. Mater. Manuf. Innov.*, 6(3):207–217, 2017. (Cited on pages 1 and 63.)
- [11] Roozbeh Dehghannasiri, Dezhen Xue, Prasanna V. Balachandran, Mohammadmahdi R. Yousefi, Lori A. Dalton, Turab Lookman, and Edward R. Dougherty. Optimal experimental design for materials discovery. *Comput. Mater. Sci.*, 129:311–322, 2017. (Cited on page 1.)
- [12] A A Benzerga, J Besson, and A Pineau. Anisotropic ductile fracture: Part I: experiments. *Acta Mater.*, 52(15):4623–4638, 2004. (Cited on page 1.)
- [13] K E Puttick. Ductile fracture in metals. *Philos. Mag.*, 4(44):964–969, 1959. (Cited on page 1.)
- [14] W.M. Garrison and N.R. Moody. Ductile fracture. *J. Phys. Chem. Solids*, 48(11):1035–1074, 1987. (Cited on page 2.)
- [15] D Curran. Dynamic failure of solids. *Phys. Rep.*, 147(5-6):253–388, 1987. (Cited on page 2.)
- [16] L.C.a Davis, C.b Andres, and J.E.c Allison. Microstructure and strengthening of metal matrix composites. *Mater. Sci. Eng. A*, 249(1-2):40–45, 1998. (Cited on page 2.)
- [17] D. Hull and D. E. Rimmer. The growth of grain-boundary voids under stress. *Philos. Mag.*, 4(42):673–687, 1959. (Cited on pages 2, 29, 48, and 65.)
- [18] T G Nieh and W D Nix. A comparison of the dimple spacing on intergranular creep fracture surfaces with the slip band spacing for copper. *Scr. Metall.*, 14(3):365–368, 1980. (Cited on page 2.)
- [19] A Gysler, G Lütjering, and V Gerold. Deformation behavior of age-hardened Ti-Mo alloys. *Acta Metall.*, 22(7):901–909, 1974. (Cited on page 2.)
- [20] M A Siddiq Qidwai, Alexis C Lewis, and Andrew B Geltmacher. Using image-based computational modeling to study microstructure – yield correlations in metals. *Acta Mater.*, 57(14):4233–4247, 2009. (Cited on pages 2, 5, 30, and 65.)
- [21] A D Rollett, R A Lebensohn, M Groeber, Y Choi, and J Li. Stress hot spots in viscoplastic deformation of polycrystals. *Model. Simul. Mater. Sci. Eng.*, 18, 2010. (Cited on pages 2, 38, 46, and 57.)

- [22] Ricardo A. Lebensohn, Anand K. Kanjarla, and Philip Eisenlohr. An elasto-viscoplastic formulation based on fast Fourier transforms for the prediction of micromechanical fields in polycrystalline materials. *Int. J. Plast.*, 32-33:59–69, 2012. (Cited on pages [3](#), [14](#), [17](#), [30](#), [32](#), [33](#), [34](#), [51](#), and [76](#).)
- [23] Marc De Graef and Michael E McHenry. *Structure of materials: an introduction to crystallography, diffraction and symmetry*. Cambridge University Press, 2012. (Cited on page [6](#).)
- [24] C.N.; Wenk H.R. Kocks, U.F.; Tomé. *Texture and anisotropy: preferred orientations in polycrystals and their effect on materials properties*. Cambridge university press, 2 edition, 1998. (Cited on pages [10](#) and [35](#).)
- [25] B. Orlans-Joliet, B. Bacroix, F. Montheillet, J. H. Driver, and J. J. Jonas. Yield surfaces of b.c.c. crystals for slip on the $\{110\} \langle 111 \rangle$ and $\{112\} \langle 111 \rangle$ systems. *Acta Metall.*, 36(5):1365–1380, 1988. (Cited on pages [11](#) and [49](#).)
- [26] U F Kocks, G R Canova, and J J Jonas. Yield vectors in f.c.c. crystals. *Acta Metall.*, 31(8):1243–1252, 1983. (Cited on pages [11](#) and [49](#).)
- [27] C. Tome and U. F. Kocks. The yield surface of h.c.p. crystals. *Acta Metall.*, 33(4):603–621, 1985. (Cited on pages [11](#) and [49](#).)
- [28] G. Y. Chin and W. L. Mammel. Competition among basal, prism, and pyramidal slip modes in hcp metals. *Metall. Mater. Trans. B*, 1(2):357–361, 1970. (Cited on pages [11](#) and [49](#).)
- [29] G Sachs and Z Zur Ableitung einer Fließbedingung. Verein Deut. *Ing*, 72:734, 1928. (Cited on page [11](#).)
- [30] G I Taylor and H Quinney. The Plastic Distortion of Metals. *Philos. Trans. R. Soc. London. Ser. A, Contain. Pap. a Math. or Phys. Character*, 230:323–362, 1932. (Cited on page [11](#).)
- [31] John D Eshelby. The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proc. R. Soc. London A Math. Phys. Eng. Sci.*, 241(1226):376–396, 1957. (Cited on page [12](#).)
- [32] J. W. Hutchinson. Elastic-Plastic Behaviour of Polycrystalline Metals and Composites. *Proc. R. Soc. A Math. Phys. Eng. Sci.*, 319(1537):247–272, 1970. (Cited on page [13](#).)
- [33] A. Molinari, G. R. Canova, and S. Ahzi. A self consistent approach of the large deformation polycrystal viscoplasticity. *Acta Metall.*, 35(12):2983–2994, 1987. (Cited on page [13](#).)

- [34] R. A. Lebensohn and C. N. Tomé. A self-consistent anisotropic approach for the simulation of plastic deformation and texture development of polycrystals: Application to zirconium alloys. *Acta Metall. Mater.*, 41(9):2611–2624, 1993. (Cited on pages 13 and 51.)
- [35] R. A. Lebensohn and C. N. Tomé. A self-consistent viscoplastic model: prediction of rolling textures of anisotropic polycrystals. *Mater. Sci. Eng. A*, 175(1-2):71–82, 1994. (Cited on page 13.)
- [36] U Fred Kocks, Carlos Norberto Tomé, and H-R Wenk. *Texture and anisotropy: preferred orientations in polycrystals and their effect on materials properties*. Cambridge university press, 2000. (Cited on page 13.)
- [37] D Peirce, R J Asaro, and A Needleman. An analysis of nonuniform and localized deformation in ductile single crystals. *Acta Metall.*, 30(6):1087–1119, 1982. (Cited on page 13.)
- [38] H. Moulinec and P. Suquet. A numerical method for computing the overall response of nonlinear composites with complex microstructure. *Comput. Methods Appl. Mech. Eng.*, 157(1-2):69–94, 1998. (Cited on pages 13 and 14.)
- [39] A Prakash and R a Lebensohn. Simulation of micromechanical behavior of polycrystals: finite elements versus fast Fourier transforms. *Model. Simul. Mater. Sci. Eng.*, 17(6):064010, 2009. (Cited on page 14.)
- [40] P. Eisenlohr, M. Diehl, R. A. Lebensohn, and F. Roters. A spectral method solution to crystal elasto-viscoplasticity at finite strains. *Int. J. Plast.*, 46:37–53, 2013. (Cited on page 14.)
- [41] R. Brenner, R. A. Lebensohn, and O. Castelnau. Elastic anisotropy and yield surface estimates of polycrystals. *Int. J. Solids Struct.*, 46(16):3018–3026, 2009. (Cited on page 14.)
- [42] R A Lebensohn. {N}-site modelling of a 3{D} viscoplastic polycrystal using the Fast {F}ourier Transform. *Acta Mater.*, 49:2723–2737, 2001. (Cited on pages 14 and 17.)
- [43] Anthony D. Mandal, Sudipto; Lao, Jacky; Donegan, Sean; Rollett. Creating representative three-dimensional microstructure for complex two-phase alloys. (under preparation). 2017. (Cited on page 14.)
- [44] Tugce Ozturk, Clayton Stein, Reeru Pokharel, Christopher Hefferan, Harris Tucker, Sushant Jha, Reji John, Ricardo A Lebensohn, Peter Kenesei, Robert M

- Suter, and Anthony D Rollett. Simulation domain size requirements for elastic response of 3D polycrystalline materials. *Model. Simul. Mater. Sci. Eng.*, 24(1):015006, 2016. (Cited on page 14.)
- [45] E Voce. A practical strain-hardening function. *Metallurgia*, 51(307):219–226, 1955. (Cited on pages 16, 32, and 51.)
- [46] E Voce. The relationship between stress and strain for homogeneous deformation. *J Inst Met*, 74:537–562, 1948. (Cited on page 16.)
- [47] Philip K Chan, Wei Fan, Andreas L Prodromidis, and Salvatore J Stolfo. Distributed data mining in credit card fraud detection. *IEEE Intell. Syst. their Appl.*, 14(6):67–74, 1999. (Cited on page 19.)
- [48] Philip K Chan and Salvatore J Stolfo. Toward Scalable Learning with Non-Uniform Class and Cost Distributions: A Case Study in Credit Card Fraud Detection. In *KDD*, volume 1998, pages 164–168, 1998. (Cited on page 19.)
- [49] Travis B Murdoch and Allan S Detsky. The inevitable application of big data to health care. *Jama*, 309(13):1351–1352, 2013. (Cited on page 19.)
- [50] Andrea Mannini and Angelo Maria Sabatini. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors*, 10(2):1154–1175, 2010. (Cited on page 19.)
- [51] James Bennett, Stan Lanning, and Others. The netflix prize. In *Proc. KDD cup Work.*, volume 2007, page 35. New York, NY, USA, 2007. (Cited on page 19.)
- [52] Indranil Bose and Radha K Mahapatra. Business data mining—a machine learning perspective. *Inf. Manag.*, 39(3):211–225, 2001. (Cited on page 19.)
- [53] Marco Dorigo and Uwe Schnepf. Genetics-based machine learning and behavior-based robotics: a new synthesis. *IEEE Trans. Syst. Man. Cybern.*, 23(1):141–154, 1993. (Cited on page 19.)
- [54] Peter Stone and Manuela Veloso. Multiagent systems: A survey from a machine learning perspective. *Auton. Robots*, 8(3):345–383, 2000. (Cited on page 19.)
- [55] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, may 2015. (Cited on pages 19, 23, and 25.)
- [56] Brody Huval, Tao Wang, Sameep Tandon, Jeff Kiske, Will Song, Joel Pazhayampallil, Mykhaylo Andriluka, Pranav Rajpurkar, Toki Migimatsu, Royce Cheng-Yue, and Others. An empirical evaluation of deep learning on highway driving. *arXiv Prepr. arXiv1504.01716*, 2015. (Cited on page 19.)

- [57] Christopher (2006). Pattern recognition Bishop and machine learning. Berlin: Springer. ISBN 0-387-31073-8. *Pattern recognition and machine learning*. Springer, Berlin, 2006. (Cited on page 19.)
- [58] Ron Kohavi and Others. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada, 1995. (Cited on page 20.)
- [59] Brian J. Parker, Simon Gunter, and Justin Bedo. Stratification bias in low signal microarray studies. *BMC Bioinformatics*, 8(326):1–16, 2007. (Cited on pages 20 and 40.)
- [60] J R Quinlan. Induction of Decision Trees. *Mach. Learn.*, 1(1):81–106, 1986. (Cited on page 20.)
- [61] J Ross Quinlan. *C4. 5: programs for machine learning*. Elsevier, 2014. (Cited on page 20.)
- [62] Leo Breiman. Random Forests. *Mach. Learn.*, 45(1):5–32, 2001. (Cited on pages 21, 42, and 53.)
- [63] A Liaw and M Wiener. Classification and Regression by randomForest. *R news*, 2(December):18–22, 2002. (Cited on pages 21 and 23.)
- [64] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, and Others. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, 12(Oct):2825–2830, 2011. (Cited on pages 22, 41, 53, 68, and 69.)
- [65] Leo Breiman. Out-Of-Bag-Estimation, 1996. (Cited on pages 23, 42, 53, 65, 66, and 68.)
- [66] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.*, pages 1–9, 2012. (Cited on pages 23, 77, and 78.)
- [67] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, and Brian Kingsbury. Deep Neural Networks for Acoustic Modeling in Speech Recognition. *IEEE Signal Process. Mag.*, (November):82–97, 2012. (Cited on page 23.)
- [68] Hao Chen, Qi Dou, Lequan Yu, Jing Qin, and Pheng Ann Heng. VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images. *Neuroimage*, pages 1–9, 2017. (Cited on pages 23, 77, 78, and 79.)

- [69] Toan Duc Bui, Jitae Shin, and Taesup Moon. 3D Densely Convolutional Networks for Volumetric Segmentation. 2017. (Cited on page 23.)
- [70] Baris Kayalibay, Grady Jensen, and Patrick van der Smagt. CNN-based Segmentation of Medical Imaging Data. 2017. (Cited on pages 23 and 77.)
- [71] Nitish Srivastava, Elman Mansimov, and Ruslan Salakhutdinov. Unsupervised Learning of Video Representations using LSTMs. *Icml*, 37:2009, 2015. (Cited on page 23.)
- [72] Brian L. Decost and Elizabeth A. Holm. A computer vision approach for automated analysis and classification of microstructural image data. *Comput. Mater. Sci.*, 110:126–133, 2015. (Cited on pages 24 and 99.)
- [73] Brian L. DeCost, Harshvardhan Jain, Anthony D. Rollett, and Elizabeth A. Holm. Computer Vision and Machine Learning for Autonomous Characterization of AM Powder Feedstocks. *Jom*, 69(3):456–465, 2017. (Cited on page 24.)
- [74] Tian Xie and Jeffrey C. Grossman. Crystal Graph Convolutional Neural Networks for Accurate and Interpretable Prediction of Material Properties. 2017. (Cited on page 24.)
- [75] Aleksandr V. Fedorov and Ivan V. Shamanaev. Crystal Structure Representation for Neural Networks using Topological Approach. *Mol. Inform.*, 1600162(383):1–8, 2017. (Cited on page 24.)
- [76] A. Asgharzadeh, H. Jamshidi Aval, and S. Serajzadeh. A Study on Flow Behavior of AA5086 Over a Wide Range of Temperatures. *J. Mater. Eng. Perform.*, 25(3):1076–1084, 2016. (Cited on page 24.)
- [77] Ravindranadh Bobbili, B. Ramakrishna, V. Madhu, and A. K. Gogia. Prediction of flow stress of 7017 aluminium alloy under high strain rate compression at elevated temperatures. *Def. Technol.*, 11(1):93–98, 2015. (Cited on page 24.)
- [78] Brian W Matthews. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta (BBA)-Protein Struct.*, 405(2):442–451, 1975. (Cited on page 27.)
- [79] U Lienert, S F Li, C M Hefferan, J Lind, R M Suter, J V Bernier, N R Barton, M C Brandes, M J Mills, M P Miller, B Jakobsen, and W Pantleon. High-energy diffraction microscopy at the advanced photon source. *J. Mater.*, 63(7):70–77, 2011. (Cited on page 30.)
- [80] Andrew D Orme, Isaac Chelladurai, Travis M Rampton, David T Fullwood, Ali Khosravani, Michael P Miles, and Raja K Mishra. Insights into twinning in Mg

- AZ31 : A combined EBSD and machine learning study. *Comput. Mater. Sci.*, 124:353–363, 2016. (Cited on pages 30 and 63.)
- [81] Michael A Groeber and Michael A Jackson. DREAM.3D: A Digital Representation Environment for the Analysis of Microstructure in 3D. *Integr. Mater. Manuf. Innov.*, 3(1):5, 2014. (Cited on pages 31 and 50.)
- [82] S. P. Donegan, J. C. Tucker, A. D. Rollett, K. Barmak, and M. Groeber. Extreme value analysis of tail departure from log-normality in experimental and simulated grain size distributions. *Acta Mater.*, 61(15):5595–5604, 2013. (Cited on pages 32 and 52.)
- [83] Tony Fast and Surya R Kalidindi. Formulation and calibration of higher-order elastic localization relationships using the MKS approach. *Acta Mater.*, 59(11):4595–4605, 2011. (Cited on page 33.)
- [84] Clarence Zener. *Elasticity and anelasticity of metals*. University of Chicago press, 1948. (Cited on page 34.)
- [85] Pedro Domingos. A few useful things to know about machine learning. *Commun. ACM*, 55(10):78, 2012. (Cited on page 35.)
- [86] Henry R Piehler. Crystal-Plasticity Fundamentals. In *ASM Handb. Vol. 22A Fundam. Model. Met. Process.*, pages 232–238. ASM International, 2009. (Cited on pages 35, 37, and 49.)
- [87] K T Aust and N K Chen. Effect of orientation on the plastic deformation of aluminum bicrystals. *Acta Metall.*, 2:136–139, 1954. (Cited on page 36.)
- [88] E M Lehecky, G Palumbo, and P Lin. Improving the Weldability and Service Performance of Nickel and Iron-Based Superalloys by Grain Boundary Engineering. *Metall. Mater. Trans. A Phys. Metall. Mater. Sci.*, 29(12):3069–3079, 1998. (Cited on pages 36, 76, and 98.)
- [89] J. Luster and M. A. Morris. Compatibility of deformation in two-phase Ti-Al alloys: Dependence on microstructure and orientation relationships. *Metall. Mater. Trans. A*, 26(7):1745–1756, 1995. (Cited on page 36.)
- [90] T. R. Bieler, P. Eisenlohr, C. Zhang, H. J. Phukan, and M. A. Crimp. Grain boundaries and interfaces in slip transfer. *Curr. Opin. Solid State Mater. Sci.*, 18(4):212–226, 2014. (Cited on page 36.)
- [91] Geoffrey Ingram Taylor. Plastic strain in metals. *J. Inst. Met.*, 62:307–324, 1938. (Cited on pages 37 and 49.)

- [92] E Smith. Cleavage fracture in mild steel. *Int. J. Fract. Mech.*, 4(2):131–145, 1968. (Cited on page 38.)
- [93] C.S. Smith. Grains, Phases and Interfaces: An interpretation of Microstructure. *Trans. AIME*, 175:15–51, 1948. (Cited on page 38.)
- [94] Oliver K Johnson and Christopher A Schuh. The uncorrelated triple junction distribution function : Towards grain boundary network design. *Acta Mater.*, 61:2863–2873, 2013. (Cited on page 38.)
- [95] Georg B. Rathmayr, Anton Hohenwarter, and Reinhard Pippan. Influence of grain shape and orientation on the mechanical properties of high pressure torsion deformed nickel. *Mater. Sci. Eng. A*, 560:224–231, 2013. (Cited on page 38.)
- [96] S Rhaipu, M L H Wise, and P S Bate. Microstructural Gradients in the Superplastic Forming of Ti-6Al-4V. *Metall. Mater. Trans.*, 33(a)(January):93–100, 2002. (Cited on page 38.)
- [97] Elizabeth A Counts, William A and Braginsky, Michael V and Battaile, Corbett C and Holm. Predicting the Hall-Petch effect in fcc metals using non-local crystal plasticity. *Int. J. Plast.*, 24(7):1243–1263, 2008. (Cited on page 38.)
- [98] J A Hanley and B J McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1):29–36, 1982. (Cited on pages 39, 53, and 66.)
- [99] Ron Kohavi and Others. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada, 1995. (Cited on page 40.)
- [100] Sudhir Varma and Richard Simon. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7:1–8, 2006. (Cited on page 40.)
- [101] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. ACM. (Cited on page 41.)
- [102] Jerome H Friedman. Greedy function approximation: a gradient boosting machine. *Ann. Stat.*, pages 1189–1232, 2001. (Cited on page 41.)
- [103] Carolin Strobl, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin, and Achim Zeileis. Conditional variable importance for random forests. *BMC Bioinformatics*, 9(23):307, 2008. (Cited on pages 42 and 73.)

- [104] Habil Zare, Gholamreza Haffari, Arvind Gupta, and Ryan R Brinkman. Scoring relevancy of features based on combinatorial analysis of Lasso with application to lymphoma diagnosis. *BMC Genomics*, 14 Suppl 1(Suppl 1):S14, 2013. (Cited on pages 42, 53, 65, 69, and 73.)
- [105] Robert Tibshirani. *Regression Selection and Shrinkage via the Lasso*, 1996. (Cited on pages 42, 65, and 68.)
- [106] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least Angle Regression. *Ann. Stat.*, 32(2):407–499, 2004. (Cited on pages 42, 54, and 69.)
- [107] Habil Zare. *FeaLect: Scores Features for Feature Selection*, 2015. (Cited on pages 42, 54, and 69.)
- [108] Israel Cohen, Yiteng Huang, Jingdong Chen, and Jacob Benesty. Pearson Correlation Coefficient. In *Noise Reduct. Speech Process.*, pages 1—4. Springer, 2009. (Cited on pages 45, 65, and 67.)
- [109] H. M Hamid, M., Lyu, H., Schuessler, B. J., Wo, P. C., Zbib. Modeling and Characterization of Grain Boundaries and Slip Transmission in Dislocation Density-Based Crystal Plasticity. *Crystals*, 7(6):152, 2017. (Cited on page 46.)
- [110] B. S. Anglin, R. A. Lebensohn, and A. D. Rollett. Validation of a numerical method based on Fast Fourier Transforms for heterogeneous thermoelastic materials by comparison with analytical solutions. *Comput. Mater. Sci.*, 87:209–217, 2014. (Cited on page 46.)
- [111] D. R. Thornburg and H. R. Piehler. An analysis of constrained deformation by slip and twinning in hexagonal close packed metals and alloys. *Metall. Trans. A*, 6(8):1511–1523, 1975. (Cited on page 49.)
- [112] Tom M. Mitchell. *Machine Learning*. McGraw-Hill Science/Engineering/Math, 1997. (Cited on page 53.)
- [113] S. L. Semiatin and T. R. Bieler. Effect of texture changes on flow softening during hot working of Ti-6Al-4V. *Metall. Mater. Trans. A*, 32(7):1871–1875, 2001. (Cited on page 54.)
- [114] Kelvin Ch’Ng, Juan Carrasquilla, Roger G. Melko, and Ehsan Khatami. Machine learning phases of strongly correlated fermions. *Phys. Rev. X*, 7(3):1–9, 2017. (Cited on page 63.)
- [115] Anton O. Oliynyk, Erin Antono, Taylor D. Sparks, Leila Ghadbeigi, Michael W. Gaultois, Bryce Meredig, and Arthur Mar. High-Throughput Machine-Learning-Driven Synthesis of Full-Heusler Compounds. *Chem. Mater.*, 28(20):7324–7331,

2016. (Cited on page 63.)
- [116] Michael E Wall, Andreas Rechtsteiner, and Luis M Rocha. Singular value decomposition and principal component analysis. In *A Pract. approach to microarray data Anal.*, pages 91–109. Springer, 2003. (Cited on page 64.)
- [117] Sebastian Mika, Bernhard Scholkopf, Alex Smola, Klaus-Robert Muller, Matthias Scholz, and Gunnar Riitsch. Kernel PCA and De-Noising in Feature Spaces. In *Adv. Neural Inf. Process. Syst.*, pages 536—542, 1999. (Cited on page 64.)
- [118] Ruslan R Hinton, Geoffrey E and Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science (80-.)*, 313(July):504–507, 2006. (Cited on page 64.)
- [119] Lei Yu and Huan Liu. Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution. In *Proc. 20th Int. Conf. Mach. Learn.*, pages 856—863, 2003. (Cited on page 64.)
- [120] Isabelle Guyon and André Elisseeff. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.*, 3(3):1157–1182, 2003. (Cited on pages 64, 65, 69, and 73.)
- [121] Laurens Van Der Maaten, Eric Postma, and Jaap Van Den Herik. Dimensionality Reduction : A Comparative Review. *J. Mach. Learn. Res.*, 10(2009):66–71, 2009. (Cited on page 64.)
- [122] Krishna Rajan, Changwon Suh, and Patricio F Mendez. Principal Component Analysis and Dimensional Analysis as Materials Informatics Tools to Reduce Dimensionality in Materials Science and Engineering. *Stat. Anal. Data Min. ASA Data Sci. J.*, 1(6):361—371, 2009. (Cited on page 64.)
- [123] Ankit Agrawal, Parijat D Deshpande, Ahmet Cecen, Gautham P Basavarsu, Alok N Choudhary, and Surya R Kalidindi. Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters. *Integr. Mater. Manuf. Innov.*, 3(8):1–19, 2014. (Cited on page 64.)
- [124] Surya R Kalidindi, Stephen R Niezgod, and Ayman A Salem. Microstructure informatics using Higher-order statistics and efficient data-mining Protocols. *J. Mater.*, 63(April), 2011. (Cited on page 64.)
- [125] Partha Dey, Joe Bible, Somnath Datta, Scott Broderick, Jacek Jasinski, Mahendra Sunkara, Madhu Menon, and Krishna Rajan. Informatics-aided bandgap engineering for solar materials. *Comput. Mater. Sci.*, 83, 2014. (Cited on page 64.)

- [126] Scott R Broderick, Joseph R Nowers, Balaji Narasimhan, and Krishna Rajan. Tracking Chemical Processing Pathways in Combinatorial Polymer Libraries via Data Mining. *J. Comb. Chem.*, 12(2):270–277, 2009. (Cited on page 64.)
- [127] Yvan Saeys, Inaki Inza, and Pedro Larranaga. Gene expression A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517, 2007. (Cited on page 64.)
- [128] Feihan Lu and Eva Petkova. A comparative study of variable selection methods in the context of developing psychiatric screening instruments. *Stat. Med.*, 33(3):401–421, 2014. (Cited on page 64.)
- [129] Jörg K. Wegner, Holger Fröhlich, and Andreas Zell. Feature selection for descriptor based classification models. 1. Theory and GA-SEC algorithm. *J. Chem. Inf. Comput. Sci.*, 44(3):921–930, 2004. (Cited on page 64.)
- [130] Mark Hall and Lloyd a Smith. Feature Selection for Machine Learning : Comparing a Correlation-based Filter Approach to the Wrapper CFS : Correlation-based Feature. *Int. FLAIRS Conf.*, page 5, 1999. (Cited on pages 65 and 67.)
- [131] Ankita Mangal and Elizabeth A. Holm. Applied Machine Learning to predict stress hotspots I: Face Centered Cubic Materials. 2018. (Cited on page 65.)
- [132] Ankita Mangal and Elizabeth A. Holm. Applied Machine Learning to predict stress hotspots II: Hexagonal Close Packed Materials. 2018. (Cited on pages 65 and 74.)
- [133] Zheng Zhao, Fred Morstatter, Shashvata Sharma, Salem Alelyani, Aneeth Anand, and Huan Liu. Advancing Feature Selection Research. *ASU Featur. Sel. Repos. Arizona State Univ.*, pages 1 – 28, 2010. (Cited on page 67.)
- [134] J Pearl. *Heuristics: Intelligent search strategies for computer problem solving*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1984. (Cited on page 67.)
- [135] Isabelle Guyon. Gene selection for cancer classification using support vector machines. *Mach. Learn.*, 46(1-3):389–422, 2002. (Cited on pages 68, 69, and 70.)
- [136] Francis R Bach. Bolasso: model consistent lasso estimation through the bootstrap. In *Proc. 25th Int. Conf. Mach. Learn.*, pages 33—40, 2008. (Cited on pages 68 and 69.)
- [137] John H Sutter, Jon M and Kalivas. Comparison of forward selection, backward elimination, and generalized simulated annealing for variable selection. *Microchem. J.*, 47(1-2):60—66, 1993. (Cited on page 69.)

- [138] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 67(2):301–320, 2005. (Cited on page 69.)
- [139] Baptiste Gregorutti, Bertrand Michel, and Philippe Saint-Pierre. Correlation and variable importance in random forests. *Stat. Comput.*, pages 1–20, 2016. (Cited on pages 70 and 73.)
- [140] Laura Toloşi and Thomas Lengauer. Classification with correlated features: Unreliability of feature ranking and solutions. *Bioinformatics*, 27(14):1986–1994, 2011. (Cited on page 73.)
- [141] Carolin Strobl, Anne-Laure Boulesteix, Achim Zeileis, and Torsten Hothorn. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*, 8:25, 2007. (Cited on page 73.)
- [142] E. A. Holm, M. A. Miodownik, and A. D. Rollett. On abnormal subgrain growth and the origin of recrystallization nuclei. *Acta Mater.*, 51(9):2701–2716, 2003. (Cited on pages 76 and 98.)
- [143] E M Lehighy and G Palumbo. On the creep behaviour of grain boundary engineered nickel. *Mater. Sci. Eng. A*, 237(2):168–172, 1997. (Cited on pages 76 and 98.)
- [144] E M Lehighy, G Palumbo, P Lin, and A Brennenstuhl. Mitigating intergranular attack and growth in lead-acid battery electrodes for extended cycle and operating life. *Metall. Mater. Trans. A Phys. Metall. Mater. Sci.*, 29(1):387–396, 1998. (Cited on pages 76 and 98.)
- [145] R C Hugo and R G Hoagland. Kinetics of gallium penetration into aluminum grain boundaries - in situ TEM observations and atomistic models. *Acta Mater.*, 48(8):1949–1957, 2000. (Cited on pages 76 and 98.)
- [146] Leman Akoglu, Mary McGlohon, and Christos Faloutsos. oddball: Spotting Anomalies in Weighted Graphs. *Proc. 14th Pacific-Asia Conf. Adv. Knowl. Discov. Data Min. - Vol. Part II*, pages 410–421, 2010. (Cited on pages 76, 100, and 102.)
- [147] Leman Akoglu, Hanghang Tong, and Danai Koutra. Graph based anomaly detection and description: A survey. *Data Min. Knowl. Discov.*, 29(3):626–688, 2015. (Cited on page 76.)
- [148] J Ding, E Ma, M Asta, and R O Ritchie. Second-Nearest-Neighbor Correlations from Connection of Atomic Packing Motifs in Metallic Glasses and Liquids. *Sci*

- Rep*, 5:17429, 2015. (Cited on pages 76 and 106.)
- [149] Yaniv Taigman, Ming Yang, Marc Aurelio Ranzato, and Lior Wolf. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. (Cited on page 77.)
- [150] J Long, E Shelhamer, and T Darrell. Fully convolutional networks for semantic segmentation. *Cvpr*, pages 3431–3440, 2015. (Cited on pages 77, 78, 82, and 90.)
- [151] Bharath Hariharan, Pablo Arbeláez, Ross Girshick, and Jitendra Malik. Hypercolumns for object segmentation and fine-grained localization. In *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pages 447–456, 2015. (Cited on pages 77, 82, and 83.)
- [152] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and Jose Garcia-Rodriguez. A Review on Deep Learning Techniques Applied to Semantic Segmentation. 2017. (Cited on pages 77 and 90.)
- [153] Andrej Karpathy, George Toderici, Sachin Shetty, Tommy Leung, Rahul Sukthankar, and Li Fei-Fei. Large-scale video classification with convolutional neural networks. *Comput. Vis. Pattern Recognit. (CVPR), 2014 IEEE Conf.*, pages 1725–1732, 2014. (Cited on page 77.)
- [154] Jeff Donahue, Lisa Anne Hendricks, Marcus Rohrbach, Subhashini Venugopalan, Sergio Guadarrama, Kate Saenko, and Trevor Darrell. Long-Term Recurrent Convolutional Networks for Visual Recognition and Description. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(4):677–691, 2017. (Cited on page 77.)
- [155] Ossama Abdel-Hamid, Abdel-rahman Mohamed, Hui Jiang, Li Deng, Gerald Penn, and Dong Yu. Convolutional Neural Networks for Speech Recognition. *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, 22(10):1533–1545, 2014. (Cited on page 77.)
- [156] Alexis Conneau, Holger Schwenk, Yann Le Cun, and Loc Barrault. Very deep convolutional networks for text classification. *Proc. 15th Conf. Eur. Chapter Assoc. Comput. Linguist. Vol. 1, Long Pap.*, 1(2001):1107–1116, 2017. (Cited on page 77.)
- [157] Yoshua and others LeCun, Yann and Bengio. Convolutional networks for images, speech, and time series. *Handb. brain theory neural networks*, 3361(10):1995, 1995. (Cited on page 77.)
- [158] Andrew L. Maas, Awni Y. Hannun, and Andrew Y. Ng. Rectifier Nonlinearities Improve Neural Network Acoustic Models. *Proc. 30 th Int. Conf. Mach. Learn.*,

- 28:6, 2013. (Cited on pages 77 and 81.)
- [159] Camille Couprie, Laurent Najman, and Yann Lecun. Learning Hierarchical Features for scene labeling. *Pattern Anal. Mach. Intell. IEEE Trans.*, 35(8):1915–1929, 2013. (Cited on page 77.)
- [160] Ashwin Bhandare, Maithili Bhide, Pranav Gokhale, and Rohan Chandavarkar. Applications of Convolutional Neural Networks. *Int. J. Comput. Sci. Inf. Technol.*, 7(5):2206–2215, 2016. (Cited on page 77.)
- [161] PY Steinkraus, Dave and Buck, I and Simard. Using GPUs for Machine Learning Algorithms. In *Doc. Anal. Recognition, 2005. Proceedings. Eighth Int. Conf.*, pages 1115–1120, 2005. (Cited on page 78.)
- [162] Andrew Simonyan, Karen and Zisserman. Very deep convolutional networks for large-scale image recognition. 2014. (Cited on page 78.)
- [163] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going Deeper with Convolutions. In *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015. (Cited on page 78.)
- [164] K He, X Zhang, S Ren, J Sun Proceedings of the IEEE, and undefined 2016. Deep residual learning for image recognition. *Openaccess.Thecvf.Com*, pages 1–9. (Cited on pages 78 and 83.)
- [165] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely Connected Convolutional Networks. 2016. (Cited on page 78.)
- [166] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Nassir Navab, Joachim Hornegger, William M Wells, and Alejandro F Frangi, editors, *Med. Image Comput. Comput. Interv. – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing. (Cited on page 78.)
- [167] Aayush Bansal, Xinlei Chen, Bryan Russell, Abhinav Gupta, and Deva Ramanan. PixelNet: Towards a General Pixel-level Architecture. *arXiv Prepr. arXiv1609.06694*, 2016. (Cited on pages 78 and 82.)
- [168] Jonathan Masci, Davide Boscaini, and Pierre Vandergheynst. ShapeNet : Convolutional Neural Networks on Non-Euclidean Manifolds. 2003. (Cited on page 78.)
- [169] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-Net : Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In *Fourth Int. Conf. 3D Vis. V-Net*, pages 567–573, 2016. (Cited on pages 78 and 79.)

- [170] Maxime Oquab, Leon Bottou, Ivan Laptev, and Josef Sivic. Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks. In *Comput. Vis. Pattern Recognit. (CVPR), 2014 IEEE Conf.*, pages 1717—1724, 2014. (Cited on pages 79 and 90.)
- [171] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.*, 22(10):1345–1359, 2010. (Cited on page 79.)
- [172] Aapo Hyvärinen, Jarmo Hurri, and Patrick O Hoyer. *Natural Image Statistics-A Probabilistic Approach to Early Computational Vision.*, volume 39. 2009. (Cited on page 80.)
- [173] Erik B Dam, Martin Koch, and Martin Lillholm. Quaternions , Interpolation and Animation. Technical report, University of Copenhagen, Denmark, 1998. (Cited on page 80.)
- [174] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.*, 15:1929–1958, 2014. (Cited on page 81.)
- [175] Sergey Ioffe and Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. 2015. (Cited on page 81.)
- [176] Peter J. Huber. Robust Estimation of a Location Parameter. *Ann. Math. Stat.*, 35(1):73–101, 1964. (Cited on page 81.)
- [177] Augustus Odena, Vincent Dumoulin, and Chris Olah. Deconvolution and Checkerboard Artifacts. *Distill*, 2016. (Cited on page 82.)
- [178] Liang-chieh Chen, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 40(4):834—848, 2015. (Cited on page 82.)
- [179] Wenzhe Shi, Jose Caballero, Ferenc Husz, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 1874—1883, 2016. (Cited on page 82.)
- [180] Eli David and Nathan S. Netanyahu. DeepPainter: Painter Classification Using Deep Convolutional Autoencoders. In Alessandro E.P. Villa, Paolo Masulli, and Antonio Javier Pons Rivero, editors, *Artif. Neural Networks Mach. Learn. – ICANN 2016*, volume 9887, pages 20–28. Springer International Publishing, 2017.

- (Cited on page [83](#).)
- [181] Y Bengio, P Simard, and P Frasconi. Learning Long Term Dependencies with Gradient Descent is Difficult. *IEEE Trans. Neural Networks*, 5(2):157–166, 1994. (Cited on page [83](#).)
- [182] Sepp Hochreiter and Paolo Frasconi. Gradient Flow in Recurrent Nets: The Difficulty of Learning LongTerm Dependencies. *A F. Guid. to Dyn. Recurr. Networks*, 2009. (Cited on page [83](#).)
- [183] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alex Alemi. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. pages 4278–4284, 2016. (Cited on page [83](#).)
- [184] John Duchi, Elad Hazan, and Yoram Singer. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *J. Mach. Learn. Res.*, 12:2121–2159, 2011. (Cited on page [85](#).)
- [185] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. In *Int. Conf. Mach. Learn.*, pages 647—655, 2014. (Cited on page [90](#).)
- [186] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks ? In *Adv. Neural Inf. Process. Syst.*, pages 3320—3328, 2014. (Cited on page [90](#).)
- [187] Amr Ahmed, Kai Yu, Wei Xu, Yihong Gong, and Eric Xing. Training hierarchical feed-forward visual recognition models using transfer learning from pseudo-tasks. In *Eur. Conf. Comput. Vis.*, pages 69–82, 2008. (Cited on page [90](#).)
- [188] Adityan Rishiyur. Neural Networks with Complex and Quaternion Inputs. 2006. (Cited on page [90](#).)
- [189] Alex Kendall, Matthew Grimes, and Roberto Cipolla. PoseNet : A Convolutional Network for Real-Time 6-DOF Camera Relocalization. In *Comput. Vis. (ICCV), 2015 IEEE Int. Conf.*, pages 2938—2946, 2015. (Cited on page [90](#).)
- [190] Chase Gaudet and Anthony Maida. Deep Quaternion Networks. 2017. (Cited on page [90](#).)
- [191] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. pages 1–14, 2016. (Cited on page [90](#).)

- [192] Mathias Niepert MATHIASNIEPERT, Neclabeu Mohamed Ahmed MOHAMEDAHMED, and Neclabeu Konstantin Kutzkov KONSTANTINKUTZKOV. Learning Convolutional Neural Networks for Graphs. *Icml*, 1, 2016. (Cited on page 90.)
- [193] Mikael Henaff, Joan Bruna, and Yann LeCun. Deep Convolutional Networks on Graph-Structured Data. *arXiv*, pages 1–10, 2015. (Cited on page 90.)
- [194] Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral Networks and Locally Connected Networks on Graphs. *Iclr*, page 14, 2014. (Cited on page 90.)
- [195] Supriya Pandhre, Manish Gupta, and Vineeth N Balasubramanian. Community-based Outlier Detection for Edge-attributed Graphs. (Cited on page 90.)
- [196] Curt Allan Bronkhorst. *Fig. 3.10, Plastic deformation and crystallographic texture evolution in face-centered cubic metals*. PhD thesis, Massachusetts Institute of Technology, 1991. (Cited on page 92.)
- [197] Michael E. Nixon, Oana Cazacu, and Ricardo A. Lebensohn. Anisotropic response of high-purity alpha-titanium: Experimental characterization and constitutive modeling. *Int. J. Plast.*, 26(4):516–532, 2010. (Cited on page 95.)
- [198] B. Zeimetz, B. A. Glowacki, and J. E. Evetts. Application of percolation theory to current transfer in granular superconductors. *Eur. Phys. J. B*, 29(3):359–367, 2002. (Cited on page 98.)
- [199] Bryan W Reed and Christopher A Schuh. Grain Boundary Networks. In *Electron Backscatter Diffr. Mater. Sci.*, pages 201–214. Springer US, 2009. (Cited on page 98.)
- [200] Christopher A Schuh, Mukul Kumar, and Wayne E King. Analysis of grain boundary networks and their evolution during grain boundary engineering. *Acta Mater.*, 51:687–700, 2003. (Cited on page 98.)
- [201] M. Frary and C. A. Schuh. Nonrandom percolation behavior of grain boundary networks in high-Tc superconductors. *Appl. Phys. Lett.*, 83(18):3755–3757, 2003. (Cited on page 99.)
- [202] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. *Phys. Rev. A*, 38(1):364–374, 1988. (Cited on page 107.)
- [203] D Stauffer and A Aharony. *Introduction to Percolation Theory*. Taylor and Francis, London, 2nd edition, 1992. (Cited on page 107.)

- [204] Ilan Breskin, Jordi Soriano, Elisha Moses, and Tsvi Tlusty. Percolation in Living Neural Networks. *Phys. Rev. Lett.*, 188102(NOVEMBER):1–4, 2006. (Cited on page [107](#).)
- [205] L. C. Hsiao, R. S. Newman, S. C. Glotzer, and M. J. Solomon. Role of isostaticity and load-bearing microstructure in the elasticity of yielded colloidal gels. *Proc. Natl. Acad. Sci.*, 109(40):16029–16034, 2012. (Cited on page [107](#).)
- [206] E Guyon, S Roux, A Hansen, D Bideau, J P Troadec, and H Crapo. Non-local and non-linear problems in the mechanics of disordered systems: application to granular media and rigidity problems. *Reports Prog. Phys.*, 53(4):373–419, 1990. (Cited on page [108](#).)
- [207] Shechao Feng and Pabitra N. Sen. Percolation on elastic networks: New exponent and threshold. *Phys. Rev. Lett.*, 52(3):216–219, 1984. (Cited on page [108](#).)
- [208] M Chubynsky and M.F. Thorpe. Rigidity percolation and the chemical threshold in network glasses. *J. Optoelectron. Adv. Mater*, pages 229–240, 2002. (Cited on page [108](#).)
- [209] M. F. Thorpe, M. V. Chubynsky, D. J. Jacobs, and J. C. Phillips. Non-randomness in network glasses and rigidity. *Glas. Phys. Chem.*, 27(2):160–166, 2001. (Cited on page [108](#).)
- [210] Donald J. Jacobs and Bruce Hendrickson. An Algorithm for Two-Dimensional Rigidity Percolation: The Pebble Game. *J. Comput. Phys.*, 137(2):346–365, 1997. (Cited on page [108](#).)
- [211] C. Moukarzel and P. M. Duxbury. Stressed backbone and elasticity of random central-force systems. *Phys. Rev. Lett.*, 75(22):4055–4058, 1995. (Cited on pages [108](#) and [112](#).)
- [212] G. Laman. On graphs and rigidity of plane skeletal structures. *J. Eng. Math.*, 4(4):331–340, 1970. (Cited on page [109](#).)
- [213] Rodney J Baxter. *Exactly solved models in statistical mechanics*. Elsevier, 1982. (Cited on page [110](#).)
- [214] M. V. Chubynsky and M. F. Thorpe. Algorithms for three-dimensional rigidity analysis and a first-order percolation transition. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 76(4):1–25, 2007. (Cited on page [110](#).)
- [215] Ana Paula Appel, Deepayan Chakrabarti, Christos Faloutsos, Ravi Kumar, Jure Leskovec, and Andrew Tomkins. ShatterPlots: Fast Tools for Mining Large Graphs. *Sdm '09*, pages 802–813, 2009. (Cited on pages [110](#) and [111](#).)

- [216] Maksim Tsvetovat and Alexander Kouznetsov. *Social Network Analysis for Startups: Finding connections on the social web*. O'Reilly Media Inc., 2011. (Cited on page [110](#).)