

Adaptive Characterization of Microstructure Dataset using a Two Stage Machine Learning Approach

Arun Baskaran^a, Genevieve Kane^a, Krista Biggs^b, Robert Hull^a, Daniel
Lewis^{a,*}

^a*Dept of Material Science and Engineering, Rensselaer Polytechnic Institute, Troy NY*

^b*Dept of Material Science and Engineering, Massachusetts Institute of Technology, Boston
MA*

Abstract

One of the goals of Materials Informatics (generally) and Image Driven Machine Learning (specifically) is extraction of quantitative data from micrographs towards an efficient characterization of microstructural features. Towards this goal, we report on a new paradigm for systematic segmentation of morphological features relevant to a given microstructure using an industrially relevant titanium alloy as an example. A two stage pipeline consisting of a classification step and a segmentation step is used to process titanium microstructures containing multiple morphological features and output quantitative measurements relevant to the particular class of microstructure identified. For the classification step, a Convolutional Neural Network is trained using the Keras API, with the architecture consisting of three convolutional layers and one fully connected layer. The microstructures are classified into three labels: “lamellar”, “duplex”, and “acicular”. A material microstructure dataset of 1225 images is established, comprised of Ti-6Al-4V alloy micrographs acquired from seven different thermal processing conditions. The CNN was trained on a dataset of 1000 images and subsequently tested on a dataset of 225 images. It reported an accuracy of **93.00 ± 1.17 %**, averaged over 5 trials incorporating a random division of the total dataset into training and test sets. For the second stage of the pipeline, image processing techniques were selected specific to the classification label of the micrograph. The area fraction of equiaxed grains is extracted from bi-modal microstructures using a marker-based watershed technique, and the area fraction of the dominant α -variant is extracted from basket-weave structures using a Histogram of Oriented Gradients (HOG) method. Computational tools similar to the proof of concept pipeline demonstrated in this work can be used by engineers to better identify microstructural features that arise due to process or material variations.

Keywords : image driven machine learning, micrograph classification, image

*Corresponding author

processing of microstructures, convolutional neural network

1. Introduction

The Materials Genome Initiative[1] aims for accelerated materials discovery and design by using computational models and data science methods. The development of powerful deep learning models in the last decade, together with the availability of labeled datasets and access to open-source libraries specific to machine learning, has led to an increased application of such models to domains which were largely dependent on analytical and numerical modeling.

The pre-cursor to neural networks, the perceptron[2], could fit only linear functions, which limited its application. The introduction of the multi-layered perceptron by Rumelhart et al[3], along with the backpropagation method to train the weights for such a network, resulted in the ability to fit complex non-linear models. LeCun et al[4] used the backpropagation method to identify handwritten digits in checks, and later[5] introduced the Stochastic Gradient Descent method as part of LeNet-5, one of the pioneering convolutional neural nets, for the purpose of document recognition. However, it took till 2010s for the widespread application of CNNs, driven by improved computational power and the availability of large, labeled, datasets.

A history of evolution of CNNs in the last decade can be gleaned from the annual ImageNet Large Scale Visual Recognition Challenge(ILSVRC)[6], which has been conducted by the ImageNet project since 2010. The first CNN to win the ILSVRC, AlexNet[7], reduced the object classification error rate from 26% to 15.3%. It was also instrumental in demonstrating the importance of a CNN's depth towards its performance. **Two major milestones in the development of deep-CNNs were the reduction in the object classification error rate on the ImageNet dataset to 6.67% and 3.57% by the winners of ILSVRC in 2014 and 2015, GoogLeNet[8] and ResNet[9] respectively.**

Within the domain of material science, the last decade has seen a rise in the use of machine learning methods, and a comprehensive review of the recent such efforts can be found in [10] and [11]. Examples of applications include drawing data-based correlations in the processing-structure-property space of materials, prediction of new materials that satisfy crystal structure or property requirements, classification and characterization of microstructures, dimensionality reduction of material descriptor space, etc. Studies that apply statistical learning for material design look at the mapping between structure and properties[12] or the dependency between processing and structure[13]. Azimi et al[14] has shown that a deep-CNN can aid towards learning patterns in data at multiple scales. Xu et al[15], implement a supervised learning method to identify the key microstructural descriptors to reduce the dimensionality of the descriptor space, resulting in faster alloy designing. Chowdhary et al.[16], employed various classification techniques such as SVM, Random Forest, Nearest Neighbor, etc., to classify between dendritic microstructures and non-dendritic microstructures as well as dendrite orientation within the microstructure. Augmentation of domain expertise with techniques such as Bayesian optimization can significantly

reduce the turnaround time for material design. A few examples of application of deep learning methods to material design and characterization problems can be found in [14, 17, 18, 19, 20].

Titanium alloys are a commercially important group of alloys that find widespread use in aerospace, biomedical, and coating applications[21, 22]. At room temperature the stable crystal structure of pure Ti is HCP(α phase), and it undergoes an allotropic transformation to BCC(β phase) at temperatures greater than 882°C. Alloys in the Ti-Al-V system can be classified according to their equilibrium phases at room temperature as α -alloys, β -alloys, or dual-phase alloys. The alloy of interest in this study belongs to the latter class of alloys, and exhibits a wide range of microstructures using a few morphological features. Depending upon the specific thermo-mechanical processing conditions, these microstructures can exhibit α lamellae, acicular morphology, or a bimodal morphology consisting of equiaxed grains and α -lamellae (duplex microstructures)[23]. Thus, they are a suitable case study for image classification and feature extraction problems. An efficient technique to generate quantitative fingerprints for material microstructures will aid towards an improved mapping between processing conditions and microstructures, and thus a better understanding of process-structure-property relationships.

Previous studies have reported varying degrees of success towards classifying and segmenting material microstructures. Campbell et al.[24] implemented the watershed algorithm to segment equiaxed grains from lamellar regions in duplex titanium alloy microstructures. A Histogram of Oriented Gradients (HOG) was implemented to separate lamellar regions, exploiting the fact that they exhibit strong spatial gradients. A pixel-wise semantic segmentation of ultra-high carbon steel microstructures using a deep-convolutional neural network was reported in [20]. In [25], DeCost et al. reports a classification accuracy of greater than 80% for a support vector machine (SVM) trained on a dataset of stainless steel microstructures.

We introduce a task pipeline to demonstrate that it is possible to classify and then quantitatively extract morphological features on micrographs that exhibit a wide variety of features. The motivation of the work is to establish a paradigm for efficient quantification of morphological features in material microstructures with the help of the deep learning tool of convolutional neural network. The pipeline, which classifies the micrographs into target labels and then implements label-specific segmentation routines, ensures that only quantitative features relevant to a particular microstructure class are extracted from the post-processing routine. For example, this methodology ensures that a watershed algorithm to segment grains will not be applied to a micrograph containing a lamellar morphology. Through this work, we establish a material-microstructure dataset of 1225 images, that can be utilized for benchmarking machine learning models for multi-class classification in the domain of material science.

Section 2 of this paper describes the collection and curation of the dataset used in this work. Section 3 describes the methodology for the classification and segmentation stages of the pipeline in detail. The results from both stages of the pipeline are reported in section 4, followed by discussion in section 5. Section 6

Table 1: List of thermal processing conditions, to which the Ti-6-4 alloy samples were subjected. The training and test datasets were assimilated from micrographs obtained from these samples.

Process ID	Holding temperature ($^{\circ}\text{C}$)	Cooling method
1	1050	Water-quench
2	1050	Air-cooled
3	1050	Furnace Cooled to 950°C , then water-quench
4	1050	Furnace Cooled to 900°C , then water-quench
5	1050	Furnace Cooled to 800°C , then air-cooled
6	1050	Furnace Cooled
7	As received	As received

brings to light some limitations of the work and some challenges facing image-driven machine learning in general.

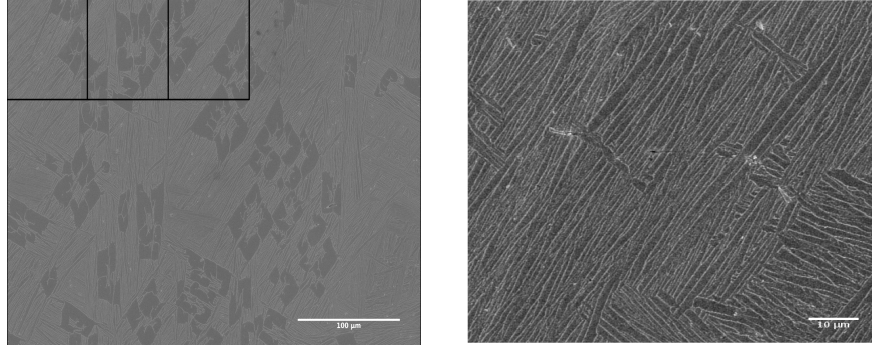
2. Dataset collection and preprocessing

2.1. Image collection methods

This work was implemented on a dataset of micrographs collected from a Ti-6%Al-4%V (Ti-6-4) alloy. Since the classifier was intended to be trained on three microstructure classes, the images were acquired from alloy samples that were subjected to a variety of thermal processing conditions in air. The authors recognize the importance of atmosphere control in the industrial processing of Ti-6-4. As this work is intended to develop representative microstructures (rather than explore heat treatments) we have made no special attempts to control the atmosphere during processing. The processing conditions are summarized in Table 1. The as-received samples were the main source of duplex microstructures. Processes 4 and 3 in the table generated the acicular microstructures, while processes 2 and 3 generated the basket-weave structures. The α -colony structures were obtained from processes 2 and 4. All specimens were imaged using scanning electron microscopy. Secondary Electron signals from the samples was utilized for generating the images. **Each image in the dataset was generated from a unique field of view to ensure that the performance of the trained network generalizes well across different fields of view in the sample.**

2.2. Dataset

Grayscale images, of size approximately $3000 \text{ pixels} \times 2000 \text{ pixels}$ were generated by SEM. The image dataset was populated by sectioning the images from source into images of size $600 \text{ pixels} \times 600 \text{ pixels}$ to ensure that the images are unambiguously annotated and the classifier is able to train on features particular to the annotated label. It was ensured that there was no spatial overlap between any two sectioned images. These sectioned images were then manually labeled as one of the three target labels. **An example instance of data augmentation described here is shown in Figure 1.**



(a) Micrograph sourced from SEM (b) Sectioned image used for training and testing purposes.

Figure 1: Visualization of the data collection method used in this work. (a) A micrograph obtained from SEM. (b) Sectioned image used in the dataset. Each black box in (a) corresponds to a sectioned image.

In this project, a dataset of 1225 images was assembled. For reference, Figure 2 provides an example image from each target label. 1000 images were randomly selected for the training process, out of which 200 were set aside as the validation set. The weights trained using this set were applied to the test dataset comprising the remaining 225 images.

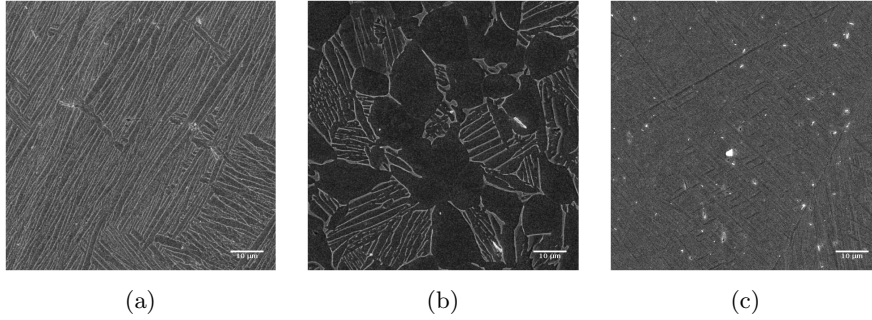


Figure 2: A representative micrograph from each target class studied in this work. (a) microstructure with a lamellar morphology, (b) a bi-modal microstructure with a mixture of equiaxed and lamellar morphology and (c) an acicular microstructure.

2.3. Data pre-processing

Pre-processing is important in a dataset of micrographs where it is likely that imaging conditions and dimensions could vary among images. For the classification step, each input image was resized into $200\text{px} \times 200\text{px}$ considering the architecture of the CNN and the computational constraints involved. Following the resizing, a Gaussian Blur with a 5×5 kernel was applied on the image.

This filter helped in reducing the inherent noise in the image and smoothing the intra-grain regions. For the segmentation step, a sequence of median smoothing, with a 5×5 kernel, followed by a contrast limited adaptive histogram equalization (CLAHE), with a 5×5 kernel, were applied to the images on the dataset. It was observed that a median smoothing retained feature edges more clearly than a mean smoothing. CLAHE was chosen over a simple histogram equalization because the former is useful in images which are non-uniformly illuminated. Figure 3 demonstrates an example of application of these filters to an original image from the dataset.

3. Methodology

3.1. Labeling

The image dataset is classified into three target labels - “lamellar”, “duplex”, and “acicular”. The images were labeled manually, and in the cases where a micrograph has multiple morphological features, the feature occupying the largest area fraction is chosen based on visual inspection.

3.2. Pipeline overview

Two main **tasks** are performed in this pipeline: classification and segmentation. Following the initial pre-processing step, images are passed to the trained classifier from which they emerge with one of the three target labels. The classifier used in this project is a convolutional neural network, which is explained in detail in the next subsection. The images along with the classification labels are passed to the post-processing stage in which label-specific segmentation is performed to extract quantitative data from the image. These label-specific tasks are explained in detail in subsection 3.4.

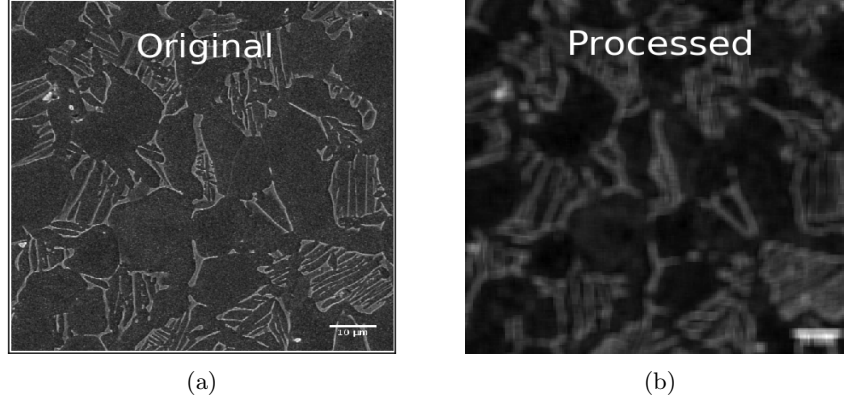
3.3. Neural network used for classification

A convolutional neural network was constructed for the purpose of classifying each image in our dataset. Classification was performed on three target labels - lamellar, duplex, and acicular microstructures. Duplex microstructures are characterized by a mixture of equiaxed regions and lamellar regions, whereas lamellae are composed of α plates. Both types of microstructures are common in Ti-alloy metallurgy.

A convolutional neural network (CNN) consists of a sequence of convolution layers, during which the network learns about the features present in the image, followed by one or many fully connected layers. When the objective of the CNN is classification, the final layer outputs the probabilities for each target label, and the label with the highest probability is selected as the predicted output. During the training phase, weights of the network are trained towards minimizing a loss function, according to a specified learning algorithm.

A convolution layer uses a filter to convolve over the input, resulting in a feature map that is dimensionally smaller than the input. Convolution with a filter F

Pre-processing before classification



Pre-processing before segmentation

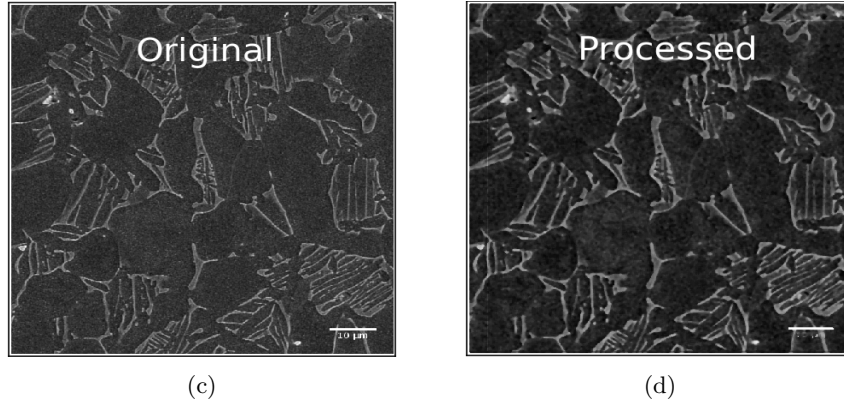


Figure 3: Application of pre-processing filters on the original image, (top) Resizing to 200px×200px followed by a Gaussian Blur with a 5×5 kernel for the classification step (bottom) Median smoothing and CLAHE, both with a 5×5 kernel, for the segmentation step.

Table 2: A dimensional representation of input and output data at every layer of the CNN. Each convolutional layer is allowed a pre-determined number of filters, which enables the network to learn the patterns inherent in the image dataset.

Layer	Input	Output	Filters	Kernel	Stride
Conv2d-1	(200,200)	(196,196)	2	5×5	(1,1)
MaxPooling-1	(196,196)	(98,98)	N/A	2×2	(2,2)
Conv2d-2	(98,98)	(94,94)	4	5×5	(1,1)
MaxPooling-2	(94,94)	(47,47)	N/A	2×2	(2,2)
Conv2d-3	(47,47)	(45,45)	12	3×3	(1,1)
Reshape	(45,45)	(24300)	N/A	N/A	N/A
Fully-Connected	(24300)	1	N/A	N/A	N/A

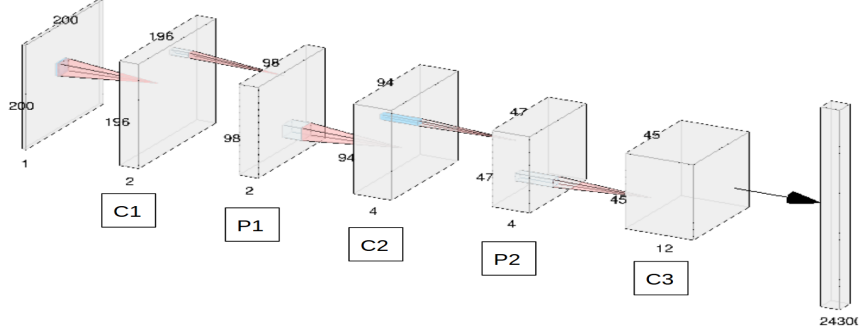


Figure 4: Visualization of architecture of the Convolutional Neural Network used in this work. The network consisted of three convolution layers, denoted by C1, C2, and C3, and one fully connected layer. A max pooling layer was used in between the convolution layers. The output to the third convolutional layer was reshaped into a one-dimensional feature vector before being passed to the fully connected layer, denoted by FC. The CNN generates a probability for each target label, and the label with the highest probability is selected as the predicted label.

of size $(k \times k)$ over the image I results in a feature map C that can be described as:

$$C(row, col) = I * F = g\left(\sum_{i=1}^k \sum_{j=1}^k I(row + i - 1, col + j - 1)F(i, j)\right) \quad (1)$$

Here, $g()$ refers to an activation function. The parameters $F(i, j)$ are referred to as the weights of the layer. The number of filters, their kernel size, and an activation function are specified for each layer. The optimum kernel size for the filters depends on the spatial scale of the features that needs to be learned by the network. While a small kernel can aid the network towards learning local features in the image, it also necessitates more number of convolution layers and hence makes the training computationally expensive. An activation function is important for introducing non-linearity in the network. In this work, a piecewise linear activation function, Rectified Linear Units (ReLU), is chosen. ReLU returns the value of the node if it is greater than zero or returns the value zero otherwise:

$$ReLU(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The weights of the filters in the convolution layers are initialized with the Xavier method[26]. This initialization draws values for the weights from a normal distribution having a mean of zero and a variance that is inversely proportional to the number of input nodes. Successive convolutional layers take the feature maps from the previous layer as input. A series of such layers enable the network to learn patterns in the image over multiple length scales. In Table 2,

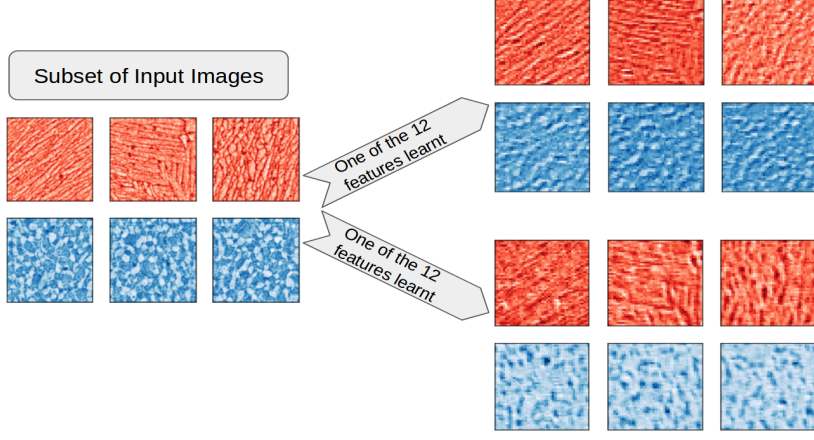


Figure 5: The CNN was provided with twelve filters to learn features present in the image. Two of the twelve features, applied over three images belonging to the lamellar class (red) and three images belonging to the bi-modal class (blue) have been shown here. While the feature at the top accentuates the oriented nature of a typical lamellar microstructure, the feature at the bottom brings out the globular pattern of equiaxed grains.

these layers are denoted by the prefix Conv2d. In this work, each convolution layer is accompanied by a max-pooling layer. Pooling layers downsample the feature maps and help in reducing the number of weights to be trained in the network. The output of a max-pooling operation P with a $(k \times k)$ kernel on an input matrix A can be described as:

$$P(row, col) = \max_{1 \leq i, j \leq k} I((row - 1) \times k + i, (col - 1) \times k + j) \quad (3)$$

They are also instrumental in making the feature maps less sensitive to the location of features in the input image. In Figure 2, the pooling layers are denoted by MaxPooling and use a kernel size of 2×2 for downsampling.

The output from the final convolutional layer is reshaped into a one-dimensional feature vector before being passed as input into a fully connected layer. Fully connected layers act as perceptrons, capable of fitting the input features to a linear function. The advantage of fitting the learned features to non-linear functions, that one can gain from having multiple fully connected layers, should be weighed against the risk of overfitting noisy patterns in the image.

The network in the current study is comprised of three convolution layers, two max-pooling layers, and one fully connected layer. The parameters defining the architecture of the CNN - number of layers, kernel size for each layer, and the number of filters for each layer - were optimized manually. The size of the training dataset, possibility of over-fitting, and the computational constraints were the important factors taken into consideration. The training hyper-parameters of the model were the learning rate, batch size, and the L1 regularization pa-

parameter, which were optimized by a grid search routine provided by the Keras library.

The kernel size and number of filters in each layer are specified in Figure 4. As an input image passes through each of these layers, the output is progressively downsampled in dimension. Twelve filters are made available to the network, over the three convolution layers, enabling it to learn twelve different features inherent in the dataset. High level features, such as edges and gradients, are learned in the first layer. Successive layers use the features from the previous layer to learn more complex shapes and abstractions present in the image.

As an example, Figure 5 shows two of the twelve features at the end of the third convolution layer, applied over three images belonging to the lamellar class (red) and three images belonging to the bi-modal class (blue). While the feature at the top accentuates the oriented nature of a typical lamellar microstructure, the feature at the bottom brings out the globular pattern of equiaxed grains.

An objective loss function is a quantitative measure of the error between the output predicted by the network and the actual output. During the training period, the weights of the network are iteratively updated towards minimizing the total loss over the entire test dataset. For the multi-label classification performed in this work, a softmax cross-entropy function is chosen as the objective loss function. The loss, L , is analytically described below:

$$L = - \sum_i y_i \log(p_i) \quad (4)$$

$$p_i = \frac{e^{a_i}}{\sum_k e^{a_k}} \quad (5)$$

where y_i is the actual label of the input i and p_i is the softmax function of the output. The softmax function converts the output from the fully connected layer to a probability distribution of the target labels. With the goal of minimizing the loss function, the weights at each layer of the CNN are updated at each iteration via a specified algorithm. In this work, AdamOptimizer[27], which improves upon the Stochastic Gradient Descent algorithm by maintaining a per-parameter learning rate, is the algorithm used, and the network was trained for 1500 iterations. The hyper-parameters of the model are summarized in Table 3. The relatively smaller dataset makes the neural network susceptible to overfitting. Hence, some regularization measures are required. In this work, we incorporate regularization during the training process through two ways. The first is through incorporation of the penalty term, $\lambda \sum_i |w_i|$, to the loss function. Referred to as the L1 regularization, this penalty term helps in feature selection by eliminating the weights connecting inconsequential features, thereby minimizing the overfitting of the model by stray features. In addition to the penalty term, an implicit early stopping rule [28] is incorporated into the training process. This rule terminates the training when the score on the validation set stops improving as compared to the previous timestep, and is a reasonable identification of the point in the training process when any further improvement in accuracy on the training data will likely not generalize well to the test data.

Table 3: Hyperparameters used during the training process.

Hyper-parameter	Description	Value
Learning Rate	Used for AdamOptimizer	0.001
Batch Size	Number of images used for optimization during every iteration	200
Max Iterations	Maximum number of iterations	1500
λ	L1 Regularization Parameter	0.001
Early Stopping criteria	Criteria for stopping the training before max_iter	$(Val_{acc}^t - Val_{acc}^{t-1}) \leq 0$

The project was implemented in Python 3, with the help of the following libraries: Numpy[29], Matplotlib[30], OpenCV[31], and Keras[32].

3.4. Algorithms used for segmentation

For the second stage of the pipeline, relevant quantitative features were extracted from images classified as either duplex or lamellar. For the feature extraction task of the pipeline, two different segmentation techniques were tested. In the first, equiaxed regions were segmented from the lamellae regions in duplex microstructures and in the second, the dominant variant was segmented from a micrograph representing a lamellar microstructure.

A marker-based topographical watershed algorithm[33] is implemented for segmenting the equiaxed regions from the lamellar regions. Spatial markers are identified corresponding to the local minimas in the gradient function. The watershed algorithm treats the image as a topological surface, with peaks and troughs corresponding to high and low pixel intensity values respectively. Regions are segmented around the markers, and boundaries between two regions are drawn with the help of intervening peaks. A more detailed description of how a topographic watershed algorithm works can be gleaned from OpenCV’s official documentation of the watershed function[31]. Subsequently, a marker-based watershed function was implemented to segment the equiaxed regions from lamellar regions.

For the second kind of segmentation, a Sobel operator from openCV was implemented to extract gradients in x (G_x) and y (G_y) directions. **The magnitude and angle of gradient at every pixel location were calculated as follows:**

$$\text{Mag} = (\sqrt{G_x^2 + G_y^2}) \quad (6)$$

$$\text{Angle} = (\tan^{-1} \frac{(G_y)}{(G_x)}) \quad (7)$$

A histogram of oriented gradients (HOG) was extracted. Figure 6 shows a typical HOG for lamellar microstructures.

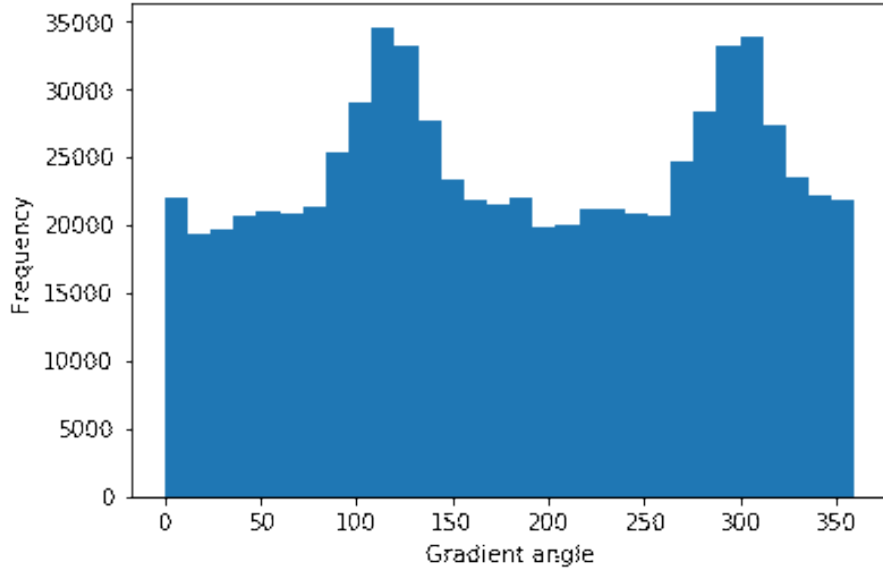


Figure 6: A characteristic Histogram of Oriented Gradients for a lamellar microstructure. The distinct peaks on the histogram represents the pixels corresponding to the dominant α variant present in the microstructure.

Due to the presence of preferentially oriented gradients, the HOG for a lamellar microstructure is likely to contain sharp peaks that represent systematic contrast variations perpendicular to the lamellar growth direction. This fact was exploited to isolate pixels corresponding to the two highest peaks on the image’s HOG. It should be noted that the use of HOG for segmenting lamellar regions in titanium alloy microstructures was reported in [24]. This concept is expanded in the current work to predict morphological features such as macroscopic variant selection. Finally, a kernel-based mean smoothing operation, similar to the procedure described in Section 2.4, was implemented on the marker for the purpose of visualization of the segmented regions.

4. Results

The performance of the trained network was quantified by the fraction of images in the test set that were classified accurately:

$$\text{Model_score} = \sum_{i=1}^M \delta_i / M \quad (8)$$

where δ_i equals 1 if the i -th image in the test dataset was classified correctly, or zero otherwise. M here refers to the size of the test dataset. The network scored consistently above 90% on test datasets of 225 images. Over five cycles

of sequentially dividing the total dataset into training and testing sets, training the network, and evaluating the trained network on the test set, an average score of 93.00 ± 1.17 % was obtained with the best score of 94.67%.confusion matrix, which can be seen in Figure 9, was generated to better visualize the network’s performance. Each row in the matrix represents the distribution of images belonging to a target label in the actual test dataset. Each column in the matrix represents the distribution of images that were classified as a particular label by the trained network. The diagonal entries denote the instances that were correctly predicted by the network, whereas entries in (Row, Column) denote the number of images belonging to the label specified by “Row” that were incorrectly classified as the label specified by “Column”. the The errors were divided uniformly across the three target labels, i.e. no single label was significantly misclassified.

Figures 7 and 8 show the results of segmentation algorithms implemented on a lamellar microstructure and a duplex microstructure respectively. The segmentation algorithm based on HOG isolates the pixels corresponding to the two highest peaks on the image’s HOG. A binary mask is created for the image, on which these isolated pixels are activated and the remaining pixels are defined to have a value of zero. The dominant α -variant in the micrograph can be visualized by superimposing this mask over the original image, as seen in Figure 7. The area fraction of the dominant α -variant is obtained by calculating the number of pixels in the mask having a value of one. Similarly for the bi-modal microstructures, a binary mask is created to isolate the pixels lying within the equiaxed grains. The fraction of the activated pixels on the mask is used as the metric for quantification of area fraction of equiaxed grains. The contours of the isolated grains are imposed on the original image for visualization of the segmentation. The complete workflow is summarized in Figure 9.

5. Discussion

The data augmentation method and the architecture of the network were tailored appropriate to the material dataset in hand. The spatial resolution in the original images was the primary factor taken into consideration before adopting the sectioning size. Among the pre-processing steps applied at various stages of the pipeline, the classification results were sensitive to the kernel size of Gaussian Blur and the magnitude of resizing. A larger cropped image resulted in increased time for training per image and fewer images in the dataset, whereas resizing to a smaller cropped image resulted in poorer classification performance. In contrast to this dataset, one containing images of lower magnification could be subjected to a finer sectioning size whereas a coarser sectioning size could be more appropriate for a dataset containing images of a higher magnification. Median smoothing was observed to be preferable to mean smoothing, because it accentuated the edges more clearly than mean smoothing. CLAHE was included as a precautionary measure and the segmentation algorithm for the images in the current material dataset were not particularly sensitive to this filter. As an input image is passed through multiple convolution layers, the features learned

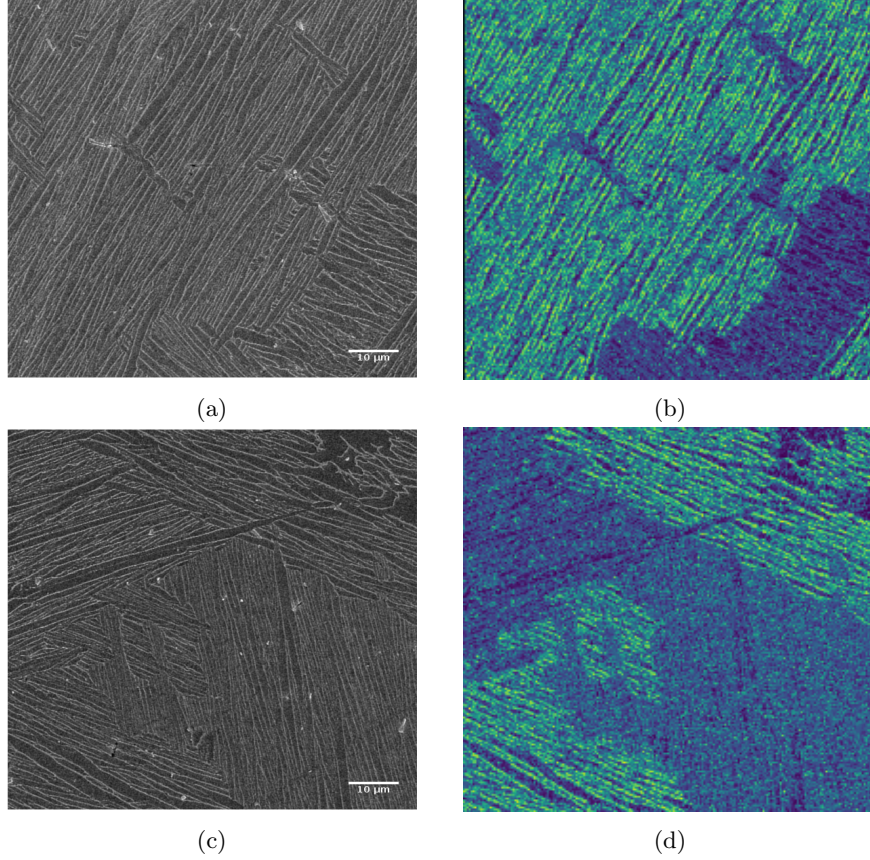


Figure 7: Segregation of the most dominant α -variant in a lamellar microstructure. Superimposing a mask of segregated pixels over the original image (a and c) results in the visualization of the dominant α -variant (denoted by the darker pixels on subfigures b and d). Such a segmentation technique can help towards a quantification of texture inherent in the micrograph.

at each layer becomes increasingly local and complex. While a manual optimization method was employed for identifying the appropriate architecture for this dataset, readers are referred to upcoming neural-network synthesis methods such as [34] and [35] for generating neural networks for complex morphologies. The number of convolution layers and the number of filters that each layer was allowed to learn were the two main parameters of the network that influenced its training performance. Too few filters resulted in the network not being able to learn the important features, whereas too many filters and an increased number of convolution layers resulted in a higher probability of overfitting.

The segmentation algorithms demonstrated here could enable a quantification of lamellae-induced texture inherent in the micrograph. Larger the area fraction occupied by the dominant α -variant, the more likely that there is a lamellae-induced texture in the micrograph. Similarly, the area fraction of

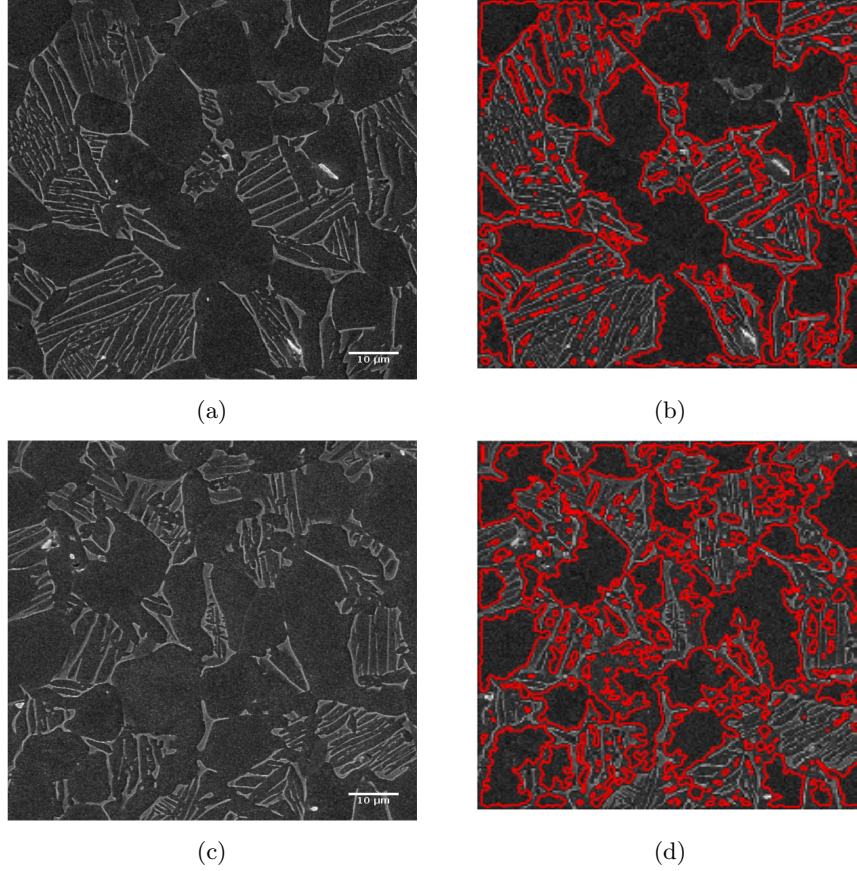


Figure 8: Segmentation of equiaxed regions in duplex microstructures. (a and c) Original images. (b and d) Segregated images.

equiaxed grains in bi-modal microstructures, Figure 8, could enable the mapping of thermo-mechanical processing conditions to the extent of globularization undergone by the material. By integrating this workflow with experimental techniques for phase analysis such as Electron Beam Scattering Diffraction (EBSD) or Energy Dispersive Spectroscopy (EDS), a better understanding of the process-structure-property relationships in the material could be explored.

This study facilitates automatic selection of the image processing method that is relevant and most useful to the given microstructure. Since the class of alloys studied here exhibits a wide variety of morphologies, multiple image processing algorithms are required to develop a quantitative fingerprint of the microstructure, and any particular algorithm might not be suitable for every microstructure. For example, if a micrograph contains a grain exhibiting lamellar morphology, information from a post-processing routine built for segmenting equiaxed grains will not provide a meaningful quantitative result for this partic-

ular morphology. An efficient method to relate microstructures to quantitative data, as shown in Figure 9, can contribute towards a robust mapping of processing conditions to properties and thereby towards better alloy design.

6. Challenges in IDML and Limitations of the Present Work

Quantitative image analysis in metallography has not yet fully realized the benefit of image driven machine learning. This is in part due to the way in which microstructure data is collected and used. It is common practice in microscopy to collect overview and detail images that identify characteristic features present throughout the structure. If statistical measurements are required, one strategy to ensure accurate sampling might be to estimate variances and compute a confidence interval from a small set of images. Recommendations provided as part of ASTM standards can also be found, however they are general in nature and leave all decision making to the analyst[36, 37]. Although no specific guidance is provided the metallographer might be instructed to collect a “suitable” number of images. The multi-step classification strategy here provides a way to manage a large collection of images that could assist the analyst in identifying and processing images more quickly than by visual examination. In this work we choose to limit the image types to SEM images where the illumination conditions and resolution are more well controlled.

It is noted that the architecture used for the neural network is designed for this material dataset, and is likely required to be fine-tuned for a dataset that contains a different number and type of features and classification targets. The relatively smaller training datasets make the domain currently not optimum for deep learning methods. However, techniques such as convolutional neural networks help towards a better interpretation of the machine learning results, despite not giving a significant improvement over conventional classifiers such as SVM or Random Forest Trees. Future innovations in the field of deep learning, such as transfer learning and neural-network synthesis, will aid towards making the application of deep learning to material science datasets ubiquitous.

7. Conclusions and Future Scope

We have demonstrated a paradigm for efficient quantitative characterization of a microstructure dataset containing different morphological features by leveraging the power of deep learning tools and image segmentation algorithms. A database of 1225 Ti-6Al-4V micrographs is utilized for this purpose.

Based on our work and observations, we put forth the following conclusions:

- A material microstructure dataset relevant to Ti alloys has been established that can be used as a benchmark for multi-class classification testing within the domain of material science.
- The CNN, built up of three convolution layers and one fully connected layer, is trained on a dataset of 1000 images. The trained network produces an accuracy of 93.00 ± 1.17 % on a testing set of 225 images.

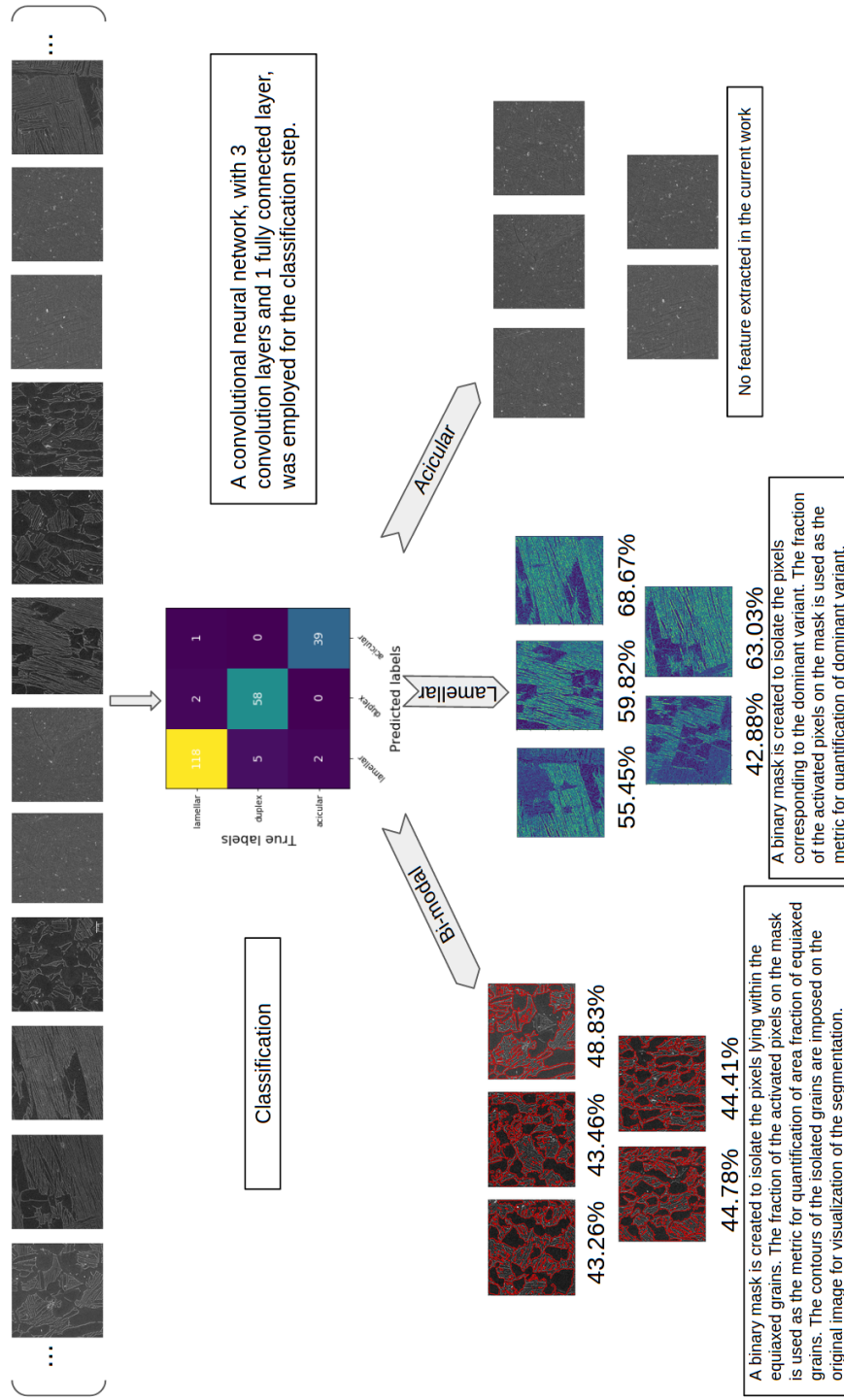


Figure 9: Overview of the pipeline implemented in this work. The first stage of the pipeline is a convolutional neural network, which classifies images into duplex (first cluster), lamellar(second cluster), and acicular(third cluster). The second stage of the pipeline implements label-specific image segmentation tasks to extract quantitative data that is relevant to the microstructure that is present in the image.

- The segmentation algorithm for lamellar morphologies extracts the area fraction of the most dominant α variant by segmenting the pixels occupying the most dominant orientations in the micrograph’s HOG. The segmentation algorithm for duplex morphologies extracts the area fraction of equiaxed grains by implementing a marker-based watershed algorithm on the micrograph.
- This approach ensures that only features relevant to the microstructure are quantified during post-processing. This results in a reduction of the time required to develop a quantitative metric for a given micrograph. It should be noted that this work is a case-study in “smart segmentation” of micrographs, and morphology-specific segmentation algorithms were selected. Based on knowledge about the specific material class and their possible morphologies, other suitable segmentation algorithms could be selected in other material systems.

8. Acknowledgement

This research is sponsored through a grant from the National Science Foundation, Award CMMI-1729336, DMREF: Adaptive control of Microstructure from the Microscale to the Macroscale. **The experimental work required for generating the material dataset was performed with the aid of clean room and characterization facilities within the Center for Materials, Devices and Integrated Systems at Rensselaer Polytechnic Institute. One of the authors (K. Biggs) completed image acquisition for this work as part an undergraduate research assignment while at Rensselaer.**

9. Data Availability Statement

The raw data and source codes required to reproduce these findings are available to download from the following repository:
<https://github.com/ArunBaskaran/Image-Driven-Machine-Learning-Approach-for-Microstructure-Classification-and-Segmentation-Ti-6Al-4V/>. A pre-trained CNN, which produced the best performance among the training cycles, has also been made available.

References

- [1] Materials genome initiative for global competitiveness.. [washington, d.c.]: Executive office of the president, national science and technology council, 2011, 2011.
- [2] F.Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 1958.
- [3] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(9), 1986.

- [4] Yann LeCun, Bernhard E. Boser, John S. Denker, Donnie Henderson, R. E. Howard, Wayne E. Hubbard, and Lawrence D. Jackel. Handwritten digit recognition with a back-propagation network. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems 2*, pages 396–404. Morgan-Kaufmann, 1990.
- [5] Yann Lecun, Leon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [6] <http://image-net.org/challenges/lsvc/>.
- [7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 25(10), 2012.
- [8] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, June 2015.
- [9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, June 2016.
- [10] Rampi Ramprasad, Rohit Batra, Ghanshyam Pilania, Arun Mannodi-Kanakkithodi, and Chiho Kim. Machine learning in materials informatics: recent applications and prospects. *Nature Computational Materials*, 54, 2017.
- [11] Frederic E. Bock, Roland C. Aydin, Christian J. Cyron, Norbert Huber, Surya R. Kalidindi, and Benjamin Klusemann. A review of the application of machine learning and data mining approaches in continuum materials mechanics. *Frontiers in Materials*, 6, May 2019.
- [12] Prasanna V. Balachandran. Machine learning guided design of functional materials with targeted properties. *Computational Materials Science*, 164:82–90, 2019.
- [13] Liu YiHao, HU ZiHeng, Suo ZhiGuang, Hu LianZhe, Feng LingYan, Gong XiuQing, and Liu Yi & Zhang JinCang. High-throughput experiments facilitate materials innovation:a review. *SCIENCE CHINA Technological Sciences*, 62(4):521–545, 2019.
- [14] Seyed Majid Azimi, Dominik Britz, Michael Engstler, Mario Fritz, and Frank Mcklich. Advanced steel microstructural classification by deep learning methods. *Scientific Reports*, 8:2128, 2018.

- [15] Hongyi Xu, Ruoqian Lu, Alok Choudhary, and Wei Chen. A machine learning-based design representation method for designing heterogeneous microstructures. *Journal of Mechanical Design*, 137, 2015.
- [16] Aritra Chowdhury, Elizabeth Kautz, Blent Yener, and Daniel Lewis. Image driven machine learning methods for microstructure recognition. *Computational Materials Science*, 123:176–187, 2016.
- [17] Ruijin Cang and Max Yi Ren. Deep network-based feature extraction and reconstruction of complex material microstructures. *ASME 2016 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, August 21-24, 2016 2016.
- [18] Nicholas Lubbers, Turab Lookman, and Kipton Barros. Inferring low-dimensional microstructure representations using convolutional neural networks. *Physical Review E*, 96, 2017.
- [19] Yanan Zhu, Qi Ouyang, and Youdong Mao. A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy. *BMC Bioinformatics*, 18(348), 2017.
- [20] Brian L. DeCost, Toby Francis, and Elizabeth A. Holm. High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel. *arxiv*, 2018.
- [21] Dipankar Banerjee and J.C. Williams. Perspectives on titanium science and technology. *Acta Materialia*, 61:844–879, 2013.
- [22] Robert Pederson. *Microstructure and Phase Transformation of Ti-6Al-4V*. PhD thesis, Lule University of Technology, 2002.
- [23] Genevieve A Kane, M. David Frey, and Robert Hull. Influence of controlled cooling rates during thermal processing of ti-6 al-4 v alloys using in-situ scanning electron microscopy. In *Submitted to MRS Advances*, 2019.
- [24] Andrew Campbell, Paul Murray, Evgenia Yakushina, Stephen Marshall, and William Ion. New methods for automatic quantification of microstructural features using digital image processing. *Materials and Design*, 141:395–406, 2018.
- [25] Brian L. DeCost and Elizabeth A. Holm. A computer vision approach for automated analysis and classification of microstructural image data. *Computational Materials Science*, 110:126133, 2015.
- [26] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In Yee Whye Teh and Mike Titterton, editors, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*, pages 249–256, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR.

- [27] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv:1412.6980v9*, 2017.
- [28] Lutz Prechelt. Early stopping-but when? In *Neural Networks: Tricks of the Trade, This Book is an Outgrowth of a 1996 NIPS Workshop*, pages 55–69, London, UK, UK, 1998. Springer-Verlag.
- [29] <https://numpy.org>.
- [30] <https://matplotlib.org>.
- [31] <https://opencv.org>.
- [32] <https://keras.io>.
- [33] Serge Beucher and F Meyer. Segmentation: The watershed transformation. mathematical morphology in image processing. *Optical Engineering*, 34:433–481, 01 1993.
- [34] Barret Zoph and Quoc V. Le. Neural architecture search with reinforcement learning. *CoRR*, abs/1611.01578, 2016.
- [35] Esteban Real, Sherry Moore, Andrew Selle, Saurabh Saxena, Yutaka Leon Suematsu, Jie Tan, Quoc V. Le, and Alexey Kurakin. Large-scale evolution of image classifiers. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML’17, pages 2902–2911. JMLR.org, 2017.
- [36] ASTM Standard E1382-97. Standard test methods for determining average grain size using semiautomatic and automatic image analysis. Technical report, ASTM International, West Conshohocken, PA, 2015.
- [37] ASTM Standard E3-11. Standard guide for preparation of metallographic specimens. Technical report, ASTM International, West Conshohocken, PA, 2017.