Article

Machine Learning Based Methods for Obtaining Correlations between Microstructures and Thermal Stresses

Akshay Bhutada ¹, Sunni Kumar ¹, Dayalan Gunasegaram ² and Alankar Alankar ^{1,*}

¹ Department of Mechanical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India; akshaybhutada@iitb.ac.in (A.B.); skbarnwal@iitb.ac.in (S.K.)

² The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Research Way, Clayton, VIC 3168, Australia; Dayalan.Gunasegaram@csiro.au

* Correspondence: alankar.alankar@iitb.ac.in

Abstract: The microstructure–property relationship is critical for parts made using the emerging additive manufacturing process where highly localized cooling rates bestow spatially varying microstructures in the material. Typically, large temperature gradients during the build stage are known to result in significant thermally induced residual stresses in parts made using the process. Such stresses are influenced by the underlying local microstructures. Given the extensive range of variations in microstructures, it is useful to have an efficient method that can detect and quantify cause and effect. In this work, an efficient workflow within the machine learning (ML) framework for establishing microstructure–thermal stress correlations is presented. While synthetic microstructures and simulated properties were used for demonstration, the methodology may equally be applied to actual microstructures and associated measured properties. The dataset for ML consisted of images of synthetic microstructures along with thermal stress tensor fields simulated using a finite element (FE) model. The FE model considered various grain morphologies, crystallographic orientations, anisotropic elasticity and anisotropic thermal expansion. The overall workflow was divided into two parts. In the first part, image classification and clustering were performed for a sanity test of data. Accuracies of 97.33% and 99.83% were achieved using the ML based method of classification and clustering, respectively. In the second part of the work, convolution neural network model (CNN) was used to correlate the microstructures against various components and measures of stress. The target vectors of stresses consisted of individual components of stress tensor, principal stresses and hydrostatic stress. The model was able to show a consistent correlation between various morphologies and components of thermal stress. The overall predictions by the model for all the microstructures resulted into $R^2 \approx 0.96$ for all the stresses. Such a correlation may be used for finding a range of microstructures associated with lower amounts of thermally induced stresses. This would allow the choice of suitable process parameters that can ensure that the desired microstructures are obtained, provided the relationship between those parameters and microstructures are also known.

Keywords: material design; data-enabled predictions; materials informatics; additive manufacturing; ICME

Citation: Bhutada, A.; Kumar, S.; Gunasegaram, D.; Alankar, A. Machine Learning Based Methods for Obtaining Correlations between Microstructures and Thermal Stresses. *Metals* **2021**, *11*, 1167. <https://doi.org/10.3390/met11081167>

Academic Editor: Chris Aldrich

Received: 3 June 2021

Accepted: 16 July 2021

Published: 22 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The need to understand microstructures and establish their relationship with properties of the parent material has been a major driver of Materials Science and Engineering research. Characteristic material information is hidden over multiple length scales and thus extracting this information and creating an interpretable representation is a key challenge. Thus, for establishing robust cause-and-effect relations for material design it is natural to understand a given dataset and to categorize it. An important aspect of this exercise is automation and performing the analyses without any human bias. Machine learning (ML) tools are ideal in such cases [1].

The close relationship between process parameters, microstructure and properties is a matter of interest in different manufacturing processes such as forming, casting, welding, sintering, and additive manufacturing, etc. [2,3]. It is of greater significance in the emerging field of additive manufacturing (AM) particularly because the process results in parts that have spatially varying properties owing to the highly localized cooling rates which determine local microstructures [3]. Therefore the microstructure and the properties on AM parts must be described using increased resolution. In other words, bulk properties, e.g., those obtained using traditional tensile tests may not be sufficiently representative of the overall properties exhibited by the material. Thus, an efficient methodology that can extract and quantify the extensive range of microstructure–property or microstructure–response relationships is required. ML is used for identification of patterns in high dimensional data and for finding the relationship between input (features) and outputs (labels). Thus, ML based reduced order models can be applied to establish process–structure–property (PSP) [4–6] linkages that turn into a mapping similar to Ashby’s map [7].

One of the earliest examples showing application of neural networks (NN)—a subset of artificial intelligence (AI) based method for material behavior—was reported by Ghaboussi et al. [8]. These researchers used NN for modeling the non-linear stress–strain behavior of composites during monotonic and cyclic loading. Haj-Ali et al. [9] applied artificial neural network (ANN) based models for addressing nonlinear micromechanical behavior and damage in heterogeneous composites. The ANN model was trained based on data generated by FE simulations of composites. Recently, Deep Learning (DL) approaches have emerged as the preferred tool for solving the problem of automatic feature identification from a wide range of possible features followed by regression [10]. Computer vision (e.g., image segmentation, image detection, and face recognition) is only one of the application domains where these approaches have been effective. In terms of learning the embedded correlations in an aggregated dataset, the DL approach greatly outperforms conventional ML approaches [11]. Within microstructure analyses, applications of DL are widespread. Liu et al. [12] presented an application of deep CNN for autoindexing during imaging. Yang et al. [13] demonstrated that the DL approaches can be employed for establishing the localization relationship between the structure and property of a three dimensional composites by learning the important features of the structures. Cecen et al. [14] showed that 3D CNN can learn salient features from material microstructures that in turn lead to good predictions of an effective property of interest. Alqahtani et al. [15] demonstrated deep CNN to predict porous media properties. The authors used grayscale micro-CT images as the input and computed porous media properties like porosity, coordination number and average pool size as the output. Rong et al. [16] determined thermal conductivity of the composites from cross-section images predicted using DL methods. Steinmetz et al. [17] demonstrated workflow for classifying synthetic microstructures generated using the Phase-Field method [18]. ML algorithms are capable of learning the linear and nonlinear relationships between the property of a material and the related process parameters [19–21].

ML algorithms are helping AM practitioners improve product quality, optimize the manufacturing process, and reduce costs [22]. Popova et al. [3] have demonstrated an efficient data-based methodology correlating AM process parameters and synthetically created microstructures using kinetic Monte Carlo (kMC). General Electric (GE) reported [23] that it used a supervised ML model to predict surface residual stresses on the hard-to-inspect internal as-built surfaces of AM parts by correlating it with microscopic surface roughness. The results were obtained in a matter of seconds in contrast to several hours required by FE simulations. Large temperature gradients experienced by an AM part during its build stages give rise to significant thermally-induced residual stresses on the parts. Residual stresses are a major concern, as they have implications for post-build distortion and low cycle fatigue performance. It is noteworthy that for AM parts, these residual stresses are rather complicated since AM is used for complicated geometries. An FE model that accounts for all the spatial variations in properties would be computationally expensive

owing to the small mesh sizes required to support the increased resolution [24]. Reduced order models can however significantly reduce the computational overhead. A detailed description of application of ML approach useful for AM technology is reported in the recent work of Wang et al. [25].

The objective of the present work was to develop an ML based reduced order model to establish correlations between microstructures and thermal stresses. Synthetically generated 2D microstructures of various grain morphologies were used with anisotropic elasticity and thermal expansion coefficients assigned to them. Columnar grain structures found in AM parts were included. Image classification algorithm was applied using CNN to classify [26–28] six different types of microstructures. A VGG16 CNN model [29] was used for extracting features from images. Principal component analysis (PCA) [30–33] was used to transform the features and also to reduce the dimensionality of features. *k*-Means clustering [34,35] was applied on these transformed features to partition the data into *k* pre-defined clusters. In the second part of the work, an NN regression model was applied to find the correlation between the microstructures and the state of thermal stress. Finally the CNN model was evaluated using multiple coefficient of determination R^2 and Root Mean Square Error (RMSE).

2. Methodology

2.1. Dataset

The dataset used for this work consisted of images of six different types of microstructures that were generated using an open source program NEPER [36]. Using NEPER, synthetic microstructures can be generated that show similar statistics and features as observed in the experiments. The output data can be generated in formats suitable for many commercial and open-source FE programs. These six types of microstructures represented six different morphologies, namely columnar grain structure with aspect ratios 1:2, 1:4, 1:6, 1:8, Voronoi based tessellation and grain growth based tessellation. The grain growth microstructure was produced based on a grain growth algorithm applied in addition to Voronoi tessellation. In total, 500 randomizations of each type of aforementioned microstructures were generated. Thus, a dataset containing a total of 3000 microstructures was created. Each microstructure was created using 120 grains and each grain was assigned a random crystallographic orientation following Bunge's convention of Euler angles [37]. Figure 1 shows representative microstructures where various colors of the grains represent crystallographic orientations. Anisotropic elastic constants in the form of a 4th rank elasticity tensor C_{ijkl} were supplied in the original basis, where i, j, k, l vary from 1 to 3. Since, in the present work the focus was on thermal stresses, elasticity constants representing very high elastic anisotropy given by monoclinic zirconia (m-zirconia) were used. Following the theory of linear elasticity, stress in an individual grain is given as

$$\sigma_{ij} = C'_{ijkl} \epsilon_{kl} \quad (1)$$

where C'_{ijkl} is the rotated tensor of elasticity constants give by

$$C'_{pqrs} = R_{pi} R_{qj} C_{ijkl} R_{rk} R_{sl} \quad (2)$$

where R matrix is a rotation matrix formed using the Euler angles. Elastic constants for *m*-zirconia were taken from the work of Zhang et al. [38]. Anisotropic thermal expansion coefficients were considered. Note that it was critical to consider both anisotropic elasticity constants and anisotropic thermal expansion coefficients to account for correct thermal stresses [39]. Evolution of thermal stress was simulated using an FE model for each microstructure via applying temperature gradient across the top (1000 K) and bottom (400 K) boundaries for a simulation time of 1 s. The left and the right faces of the microstructures were fixed along the X direction and the bottom face was fixed along the Y direction. The expansion along the Y direction was possible along the Y direction at the top face

only. The FE simulations were performed using Multiphysics Object-Oriented Simulation Environment (MOOSE). MOOSE is an open-source, parallel FE framework [40].

Figure 1. Data consisting of six different types of microstructures. (a) Columnar grain structure with aspect ratio 1:2, (b) Columnar grain structure with aspect ratio 1:4, (c) Columnar grain structure with aspect ratio 1:6, (d) Columnar grain structure with aspect ratio 1:8, (e) Grain growth (gg) grain structure (f) Voronoi grain structure. Note that the grain growth structure was produced based on a grain growth algorithm applied in addition to Voronoi tessellation. This algorithm is part of the NEPER program.

Output data consisted of the fields of stress components σ_{xx} , σ_{yy} and σ_{xy} . Based on these, principal stresses σ_1 , σ_2 and hydrostatic stress σ_h were determined. The original images of the microstructures were collected in the form of *png* images of the dimensions 269 pixel $\times$ 269 pixel $\times$ 3 pixel. These microstructure images were resized into 224 pixel $\times$ 224 pixel $\times$ 3 pixel. Each pixel of an image represented a feature and carried an RGB color and thus contained three channels of information. Figure 2 shows the distribution of principal stresses σ_1 , σ_2 and hydrostatic stress σ_h across different types of microstructures. As expected, the principal stress σ_1 was not significant for most microstructures. The magnitude of principal stress σ_1 was in the range 30 to 60 MPa. The distribution of principal stress σ_2 appeared to follow a normal distribution between -1200 and -1050 MPa. The distribution of hydrostatic stress also followed a normal distribution and was between -580 and -500 MPa.

Figure 2. Distribution of (a) principal stress σ_1 (b) principal stress σ_2 and (c) hydrostatic stress σ_h across different grain structures.

2.1.1. Classification and Clustering

As mentioned earlier, the overall objective of this work was to understand the effect of microstructures on thermal stresses. For developing this correlation, it was critical to establish that the ML based methodology was able to distinguish between such structures. A CNN model was used for classification. CNN corresponds to a family of Deep Neural Network (DNN) typically used for image classification and recognition [41]. Figure 3 shows the architecture of the CNN that was used for classifying the microstructures.

Figure 3. Architecture of CNN used for classification of microstructures.

As evident from Figure 3, the CNN model consisted of an alternate stack of convolutional layers and pooling layers. The goal of the convolution layer was to extract the important features from the input image. This was achieved by convolution filters in each convolution layer [41]. In the present study, the size of the input microstructure images to the CNN model was $224 \times 224 \times 3$ where each pixel represented a feature. Note that as described earlier, in the present case each pixel carried an RGB color and thus contained three channels of information. This represented very high dimensional data and thus the dimensionality was reduced by identifying the features that had highest variance.

In the image classification process, manual labelling of images was performed for six different types of grain morphologies. Columnar grains with aspect ratios 1:2, 1:4, 1:6, 1:8, Voronoi and grain growth were labelled as 0, 1, 2, 3, 4, 5, respectively. After that, the images were randomly shuffled so that there was no bias towards a particular output. The dataset was divided into 70:15:15 for training, validation and testing. Note that the test data were

unseen by the model during regression and was used only to assess the final performance. The CNN model used is shown in Table 1. CNN consisted of a four convolution layer of channels 264, 128, 64 and 32 with kernel size of (3,3) and average pooling layer with size of (2,2). A Rectified Linear Unit (ReLU) function [41] was used as an activation function. After stacking of the convolution and average pooling layer, feature map of size (12,12,32) was achieved. This feature map was flattened to a vector of size 4608. This vector was then used as an input to the fully connected network consisting of one dense hidden layer of 32 units. Finally, output from this layer was fed to final output layer with six classes and Softmax activation function [41]. L2 regularization was used to reduce overfitting. Nesterov accelerated Adaptive Moment Estimation (NAdam) [42] optimizer was used with learning rate 0.00001. Batch size used in the model was 50 and the model was run for 500 epochs. Models were evaluated with classification accuracy and a confusion matrix. Classification accuracy is given by

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

where *TP* is true positive, *TN* is true negative, *FP* is false positive and *FN* is false negative [43,44].

Table 1. CNN for image classification. Size of input image was $224 \times 224 \times 3$ where 3 represents the three color channels of R(red), G(green) and B(blue) type.

Type	Filter Size	Number of Filters	Output Size
Convolution	3×3	264	$222 \times 222 \times 264$
Average Pooling	2×2	-	$111 \times 111 \times 264$
Convolution	3×3	128	$109 \times 109 \times 128$
Average Pooling	2×2	-	$54 \times 54 \times 128$
Convolution	3×3	64	$52 \times 52 \times 64$
Average Pooling	2×2	-	$26 \times 26 \times 64$
Convolution	3×3	32	$24 \times 24 \times 32$
Average Pooling	2×2	-	$12 \times 12 \times 32$
Flatten	-	-	4608
Dense	-	-	32
Dense	-	-	6

To further ensure that the data were not biased and was suitable for regression, image clustering was performed without labelling the input image data. Features were extracted using VGG16 [29] CNN model. VGG16 [29] is a very deep convolution model. It has 16 layers out of which 13 are convolution layers and 3 are dense layers. After the stack of convolution and the max pooling layer, a feature map of size (7,7,512) was obtained. After flattening, there were three fully connected layers. All the hidden layers used ReLU as its activation function. The model was trained using transfer learning [45]. In transfer learning, the VGG16 model was trained on some other data as the ImageNet dataset [46] in this case and then pre-trained weights were used for training the model on the actual dataset. Principal component analysis (PCA) was applied on these extracted features from VGG16 model to reduce the dimension and to better visualise the data using a 3D cluster plot.

Finally, *k*-Means clustering was applied on the transformed features from PCA. *k*-Means clustering algorithm is an iterative algorithm that splits the dataset into *k* numbers of pre-defined distinct clusters such that each data point represents only one class. It assigns the data points to various clusters such that the sum of the squared distance between the data points and the centroid of the cluster is minimum. Figure 4 summarises the methodology of image clustering.

Figure 4. Framework of methodology for clustering the images using *k*-Means clustering method.

The objective function for *k*-Means clustering is given by [41]:

$$J = \sum_{j=1}^k \sum_{i=1}^n ||x_i^j - \mu_j||^2 \quad (4)$$

In the above equation, *k* is the number of clusters, *n* is the number of cases, μ is the centroid of cluster *j*, $||x_i^j - \mu_j||$ is the Euclidean distance from the centroid of the cluster to the datapoint. The model was evaluated using an accuracy score, confusion matrix and elbow curve. The elbow method gives an intuition about the optimum number of clusters *k*. The square of the squared error (SSE) between the data points and their assigned cluster centroids is used to find *k* where SSE starts to saturate and forms an elbow.

2.1.2. Regression between Microstructures and Stresses

After ensuring the unbiased state of the data, a regression analysis was performed. The objective of the regression analysis was to establish a microstructure–thermal stress correlation in the form of an ML model. Note that various measures of thermal stresses were available as averaged over the domain from the FE simulation data. However, for our analyses, principal stresses σ_1 , σ_2 and the hydrostatic stress σ_h were chosen since many failure criteria are described as functions of these stress components. For completeness, a similar analysis using σ_{xx} , σ_{yy} and σ_{xy} is shown in Appendix A. An NN model was used to find a relationship between the microstructures and stresses. Regression analysis was performed against the above three stresses as a vector $[\sigma_1, \sigma_2, \sigma_h]$. This was useful since it assembled all necessary components required for describing failure due to thermal stresses. Images were randomly shuffled and the image dataset was divided into 70:15:15 for training, validation and test data, respectively. As noted earlier, the test data were unseen by the model during regression and was used only to test the final performance. The components of these vectors were scaled in the range of 0 and 1 and were correlated with images of microstructures as triplets. Features were extracted using a CNN model for which the architecture is shown in Table 2. The CNN model consisted of 3 convolution layers of channels 264, 128 and 64 with kernel size (3,3) and 3 average pooling layers with size (2,2). Feature map size from last average pooling layer was (26,26,64) which was flattened to a vector. The methodology of microstructure–stress regression is summarized in Figure 5.

Figure 5. Methodology of microstructure–stress regression.

Features extracted from the CNN had 43,264 dimensions. PCA was used to reduce the dimensionality. Manual search for optimum number of principal components was performed. Transformed features were then further used in the NN regression model as input. Manual hyperparameter tuning was performed for finding a suitable number of hidden layers and hidden units to select the best hyperparameters. In the NN regression model, three hidden layers were used with 2048 neurons and ReLU activation in each layer. Table 3

shows the NN regression model architecture. Dropout [47] and L2 regularization [48] were used to reduce the overfitting in the model. Dropout refers to dropping out hidden units which were randomly chosen during training in hidden layers in an NN model. Figure 6a,b show a fully connected NN and a network with dropouts, respectively.

Table 2. CNN for feature extraction from input images of size $224 \times 224 \times 3$.

Type	Filter Size	Number of Filters	Output Size
Convolution	3×3	264	$222 \times 222 \times 264$
Average Pooling	2×2	-	$111 \times 111 \times 264$
Convolution	3×3	128	$109 \times 109 \times 128$
Average Pooling	2×2	-	$54 \times 54 \times 128$
Convolution	3×3	64	$52 \times 52 \times 64$
Average Pooling	2×2	-	$26 \times 26 \times 64$
Flatten	-	-	43,264

Figure 6. (a) Fully connected NN, (b) a network applying dropout.

A regression model which uses L2 regularization is also called Ridge regression. Ridge regression adds the square of coefficient as the penalty term to the loss function. The loss function used in the NN regression model is regularized mean squared error between the true value y and the predicted value $\hat{y}$ [10]. The optimization problem was defined as follows:

$$\min \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^k w_j^2 \quad (5)$$

where $\hat{y}_i$ is an estimated stress for true value y_i , λ is regularization parameter, N is the number of images of microstructures and k is the number of weight in the model. A weight in an NN model is a trainable parameter that transforms the input data in the hidden layers. Backpropagation using the NAdam optimiser was used to update regression model parameters [42].

Table 3. NN regression model with input size 150.

Type	Hidden Units	Output Shape
Dense	2064	2064
Dropout	-	2064
Dense	2064	2064
Dropout	-	2064
Dense	2064	2064
Dropout	-	2064
Dense	3	3

The model was evaluated by multiple coefficients of determination R^2 and root mean square error (RMSE). RMSE is given by

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (6)$$

where y_i is the actual value from simulation and $\hat{y}_i$ is the predicted value from NN regression model. Coefficient of multiple determination R^2 is given by

$$R^2 = 1 - \frac{SSE}{SS_y} \quad (7)$$

where $SSE = \sum (y - \hat{y})^2$ and $SS_y = \sum (y - \bar{y})^2$. y and $\hat{y}$ are the true value and the predicted value and $\bar{y}$ is the mean of y values. The coefficient of multiple determination compared the accuracy of the model with the accuracy of multiple samples [49]. $R^2 = 1$ and $R^2 = 0$ represent perfectly fitting and poorly fitting models, respectively.

3. Results and Discussion

3.1. Classification and Clustering

The performance of the CNN model for the classification problem is shown in Figure 7a,b. Accuracy reached nearly 1.0 in approximately 100 epochs and loss sharply decreased upto 200 epoch and then gradually diminished. Classification accuracy on training, validation and test data using CNN model was 100%, 98.96% and 97.33%, respectively. Figure 7c shows a confusion matrix marking that, out of 450 test data, 439 images were correctly classified and 11 images were misclassified.

Figure 7. Variation of CNN model. (a) accuracy and (b) loss with epoch for training and validation data. (c) Confusion matrix between a known label and a predicted label for 6 classes on the test data as performed by the CNN model.

In the clustering problem, the image labels were predicted using k -Means clustering method. As pointed out earlier, known labels were used to evaluate the performance of the model. Figure 8 shows PCA analysis from which it could be concluded that 40% of variance was explained by 5 principal components and 1200 principal components were needed to explain approximately 90% variance.

As noted earlier, k -Means clustering was performed for sanity check of the data. Predicted 3D clusters of microstructures as shown in Figure 9a and the confusion matrix as shown in Figure 9b support that the model was able to identify the microstructures correctly. Out of total 3000 images 2995 images were correctly predicted by k -Means clustering. The accuracy of the model was 99.83 percent. Based on the Elbow curve as shown in Figure 9c, it was concluded that the optimum value of k was in the range 4–6 since after 6 numbers of clusters, SSE (see Equation (7)) did not decrease steeply.

Figure 8. PCA analysis showing (a) cumulative explained variance by 5 principal components, (b) number of components required to explain 95 % of the cumulative variance.

Figure 9. (a) 3D cluster plot showing predicted labels. (b) Confusion matrix between known labels and predicted labels for the 6 classes as performed by *k*-Means clustering with classification accuracy of 99.83 percent, (c) Elbow curve showing the variation of SSE as a function of number of clusters for selecting a suitable number of clusters in *k*-Means clustering.

3.2. Regression between Microstructures and Stresses

From classification and clustering results presented in Section 3.1 it was concluded that the data were suitable for regression analysis. After the feature extracted from CNN, PCA was used for dimensionality reduction. The performance of the model could be enhanced by selecting the optimum number of principal components. To find the optimum number of principal components, an NN model was trained by varying the number of principal components while all other parameters remained unchanged. Figure 10a shows the variation of R^2 and RMSE, respectively, as a function of principal components for the test data. The ideal number of principal components were selected based on maximum R^2 and minimum RMSE. As shown in the figure, with an increase of principal component, R^2 increased and RMSE decreased up to 150. After 150 principal components R^2 decreased and RMSE increased. Thus, the optimum number of principal components was found to be 150. Thus, the transformed features from PCA with dimension 150 were used for the regression model. Hyperparameter tuning was performed for hidden units and hidden layers. Figure 10b shows the variation of R^2 and RMSE as a function of number of hidden units. Model performance of NN regression increased with an increase in the number of hidden units. The optimum hidden units for best model performance was 2064 units.

As evident from Figure 10c, R^2 increased and RMSE decreased as the number of hidden layers increased from 2 to 3 but further increase in the number of hidden layers led to a decrease in R^2 and an increase in RMSE for test data. Thus, 3 hidden layers were used in the NN model.

After performing the hyperparameter tuning it was concluded that 3 hidden layers and 2064 hidden units in each layer were optimum hyperparameters for the best performance. Thus, the final NN had 3 hidden layers with 2064 hidden units in each hidden layer. The learning rate used in the final NN model was 0.00015. A batch size of 200 was used. The

dropout parameter was 0.1 and regularization $L2$ parameter was 0.00001. Figure 11a,b show the performance of the NN model quantified by regularized loss and regularized MSE, respectively, for the training and validation data. The NN model was run for a total of 2000 epochs. After 800 epochs, the training and validation loss did not decrease any further.

Figure 10. Variation of (a) R^2 and RMSE for selection of principal components, (b) R^2 and RMSE as a function of number of hidden units, (c) R^2 and RMSE as a function of number of hidden layers in NN.

Figure 12 shows the predictions by NN regression model for normalized principal stresses σ_1 , σ_2 and hydrostatic stress σ_h , respectively. The figures in the first column show the comparison for training data, the figures in the second column show the comparison for validation data and the figures in the third column show the comparisons for the test data. The test data were unseen by the model during regression and was used only to measure the final performance. The solid line in each plot represents the line of ideality. The regression model performed well for correlating microstructures to thermal stresses for all the morphologies except for the grain growth morphology. Note that as pointed out in Section 2.1, grain growth morphology was created by using a grain growth algorithm in addition to the Voronoi tessellation algorithm.

Figure 11. Variation of (a) regularized loss as a function of epochs for training and validation data (b) regularized MSE as a function of epochs for training and validation data.

R^2 and RMSE metrics for principal stresses σ_1 , σ_2 and hydrostatic stress σ_h and their average represented by $[\sigma_1, \sigma_2, \sigma_h]$ achieved from NN model are summarized in Table 4. The NN regression model was able to capture the correlation between microstructures and thermal stresses well. For principal stress σ_2 and hydrostatic stress σ_h model performance was better than principal stress σ_1 . RMSE for σ_2 and σ_h were 4.96% and 4.83% as compared to 5.73% for σ_1 .

Figure 12. NN model predictions of (a–c) σ_1 , (d–f) σ_2 and (g–i) σ_h . The figures in the first column show the comparison for training data, the figures in the second column show the comparison for validation data and the figures in the third column show the comparisons for the test data. The solid diagonal line represents the line of ideality.

Table 4. R^2 and RMSE for principal stresses and hydrostatic stress on training, validation and test data.

Target	R^2			RMSE		
	Training Data	Validation Data	Test Data	Training Data	Validation Data	Test Data
σ_1	0.9990	0.8380	0.8665	0.005	0.0613	0.0573
σ_2	0.9993	0.8581	0.8753	0.0038	0.0538	0.0496
σ_h	0.9993	0.8605	0.8735	0.0038	0.0523	0.0483
$[\sigma_1, \sigma_2, \sigma_h]$	0.9992	0.8522	0.8717	0.0042	0.0558	0.0517

Figure 13 shows the overall performance of the model assessed by making predictions using the fully trained and tested model against all the data. R^2 was approximately 0.96 for the three components of stress and RMSE was 0.82, 4.13 and 2.13 for σ_1 , σ_2 and σ_h , respectively, as shown in Figure 13a–c. In these plots representing the overall performance, the stresses have been converted back to the original scale, i.e., the one used as input. Model performance was weakly dependent on the type of microstructure. Figure 13d shows percentage residuals for different types of microstructures in relation to principal stresses

σ_1 and σ_2 , hydrostatic stress σ_h and their average $[\sigma_1, \sigma_2, \sigma_h]$. The highest % residual was recorded for grain growth based microstructure. % residuals across all the morphologies were in the range 1 to 7%. The model was able to show consistent correlations across all the morphologies for principal stresses σ_1 , σ_2 and hydrostatic stress σ_h .

In this work, we have demonstrated the application of an NN based model for establishing correlations between microstructures and thermal stresses. Training an NN model consumes a considerable time. However, once the model is fully trained the predictions can be made relatively quickly [10]. As pointed out earlier, currently thermo-elasticity based FE models, phase-field based FE models or cellular automata models are used for studying thermally induced residual stresses or temperature induced microstructure evolution. Note that FE simulations make use of constitutive equations for addressing the material behavior. In simplified FE models, anisotropic elasticity and thermal expansions are not taken into account. Thus, the residual thermal stresses are restricted to be functions of temperature gradients only [24]. If complex microstructure details are considered, the model become computationally expensive, e.g., like Phase-Field models [18]. The reduced order model approach that has been proposed here is a significant step towards adapting a multiscale approach where the material can be approximated via a microstructure sensitive reduced order NN model rather than equations that are ad hoc fitted against the experimentally observed behavior. This NN based correlation can be used in an FE model to replace the equations that are usually established based on prior knowledge. Such NN correlations can be updated as soon as new data from experiments is available. Notable is that the methodology presented here is very generic and is easily applicable for correlations such as microstructure–property where the property can be yield strength, ultimate tensile strength, hardness, fatigue life etc. Furthermore, important to note is that performing such analysis solely based on experimental data may not be feasible since it may lead to significant expenses. However, using experimental data, FE based models can be calibrated and more synthetic and realistic data can be generated as demonstrated in the present work. Such data incorporates multiphysics information embedded over multiple length-scales. In this work, anisotropic elasticity constants and thermal expansion constants have been used for determining thermal stresses based on crystallographic orientations of grains in a microstructure which makes the NN based correlation a microstructure sensitive one within the thermo-elastic regime. Such a workflow can be bridged to another ML based model where aforementioned material properties are determined as a function of composition [50]. A hierarchical model thus developed will be able to predict thermal stresses for a wide range of multicomponent alloy microstructures when used in conjunction with an FE model.

Figure 13. Cont.

Figure 13. Overall predictive performance of the NN regression model. Prediction of (a) σ_1 , (b) σ_2 and (c) σ_h based on the input of all the data to the trained and tested model. Note that in these plots representing the overall performance, the stresses have been converted back to the original scale, i.e., the one used as input for ML. (d) Residual % of NN model across various grain structures.

4. Conclusions

In this work, a generic workflow suitable for identifying and correlating the microstructure vs. thermal stress relationship has been presented. It is emphasized that a domain understanding of the various types of microstructures is critical for explaining the structure–property relations and inferences based on the relationships. A CNN model was used for classification and then regression. In the classification problem, the model was able to efficiently classify the given six classes of microstructures. For image classification, CNN model was used for feature extraction and then these features were used for classification. *k*-Means clustering method was applied by extracting features from images using a CNN model. Accuracies of 97.33% and 99.83% were achieved using the classification and clustering ML method, respectively. In the next problem a different architecture of CNN was used for regression between these six microstructures and their corresponding principal stress. Different regression algorithms were applied to correlate the microstructure against various components and measures of stress. The target vector consisted of individual components of stress tensor, principal stress and hydrostatic stress. The model was able to show a consistent correlation between various morphologies and components of thermal stress. It is emphasised that a regression between microstructure and the vector of stresses is more realistic than the regression between individual stress components since it avoids the redundant degrees of freedom between the stress components. The overall predictions by the model for all the microstructures resulted into $R^2 \approx 0.96$ for all the stresses. This workflow can be used for choosing a microstructure with the least thermal stresses. Such workflows can be employed for quantitative comparisons between experimental and model predicted internal stresses. Additionally, in conjunction with FE model for AM process, the NN model can be used as a reduced order microstructure sensitive model for predicting thermal stress on an AM part.

Author Contributions: Conceptualization, A.A.; methodology, A.A. and A.B.; software, A.B. and S.K.; validation, A.B. and A.A.; formal analysis, A.B. and S.K.; investigation, A.B., S.K., D.G. and A.A.; resources, A.A.; data curation, S.K. and A.B.; writing—original draft preparation, A.B.; writing—review and editing, A.A., D.G. and S.K.; visualization, A.B., S.K. and A.A.; supervision, A.A.; project administration, A.A.; funding acquisition, A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Aeronautics Research and Development Board (ARDB) grant number GTMAP-1952.

Data Availability Statement: The data investigated in this study will be used for further studies and thus, is not archived publicly.

Acknowledgments: A.A. would like to acknowledge project GTMAP-1952 funded by Aeronautics Research and Development Board (ARDB).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Regression Analysis Performed Based on the CNN Model Using Stress Components σ_{xx} and σ_{xy}

Regression analysis performed based on the CNN model using stress components σ_{xx} and σ_{xy} . The magnitude of σ_{yy} was very small (in the range 10^{-9} – 10^{-10} MPa). Thus, the model was not able to perform well on σ_{yy} . Figures A1 and A2 show NN model performance for σ_{xx} and σ_{xy} .

Figure A1. NN model prediction of normal stress σ_{xx} for (a) training data (b) validation data (c) test data. The solid line represents the line of ideality. Asp1:2, Asp1:4, Asp1:6, Asp1:8 and gg represents columnar grain structure of aspect ratio 1:2,1:4,1:6,1:8 and grain growth structure, respectively.

Figure A2. NN model prediction of normal stress σ_{xy} for (a) training data (b) validation data (c) test data. The solid line represents the line of ideality. Asp1:2, Asp1:4, Asp1:6, Asp1:8 and gg represents columnar grain structure of aspect ratio 1:2,1:4,1:6,1:8 and grain growth structure, respectively.

References

1. Liu, R.; Yabansu, Y.C.; Yang, Z.; Choudhary, A.N.; Kalidindi, S.R.; Agrawal, A. Context Aware Machine Learning Approaches for Modeling Elastic Localization in Three-Dimensional Composite Microstructures. *Integr. Mater. Manuf. Innov.* **2017**, *6*, 160–171. [[CrossRef](#)]
2. Maizza, G.; Caporale, A.; Polley, C.; Seitz, H. Micro-Macro Relationship between Microstructure, Porosity, Mechanical Properties, and Build Mode Parameters of a Selective-Electron-Beam-Melted Ti-6Al-4V Alloy. *Metals* **2019**, *9*, 786. [[CrossRef](#)]
3. Popova, E.; Rodgers, T.M.; Gong, X.; Cecen, A.; Madison, J.D.; Kalidindi, S.R. Process-Structure Linkages Using a Data Science Approach: Application to Simulated Additive Manufacturing Data. *Integr. Mater. Manuf. Innov.* **2017**, *6*, 54–68. [[CrossRef](#)] [[PubMed](#)]

4. Yuan, S.; Jiao, Z.; Quddus, N.; Kwon, J.S.I.; Mashuga, C.V. Developing quantitative structure–property relationship models to predict the upper flammability limit using machine learning. *Ind. Eng. Chem. Res.* **2019**, *58*, 3531–3537. [\[CrossRef\]](#)
5. Wang, Z.; Huang, S.; Wen, G.; Liu, Q.; Tang, P. Thermal conductivity prediction and structure–property relationship of CaO–SiO₂–Al₂O₃ ternary system: A combination of molecular dynamics simulations and machine learning. *J. Mol. Liq.* **2021**, *324*, 114697. [\[CrossRef\]](#)
6. Zhang, J.; Wang, P.; Gao, R.X. Deep learning-based tensile strength prediction in fused deposition modeling. *Comput. Ind.* **2019**, *107*, 11–21. [\[CrossRef\]](#)
7. Ashby, M.F.; Cebon, D. Materials selection in mechanical design. *Le J. De Phys. IV* **1993**, *3*, C7-1–C7-9. [\[CrossRef\]](#)
8. Ghaboussi, J.; Garrett, J.H.; Wu, X. Knowledge-Based Modeling of Material Behavior with Neural Networks. *J. Eng. Mech.* **1991**, *117*, 132–153. [\[CrossRef\]](#)
9. Haj-Ali, R.; Pecknold, D.A.; Ghaboussi, J.; Voyiadjis, G.Z. Simulated Micromechanical Models Using Artificial Neural Networks. *J. Eng. Mech.* **2001**, *127*, 730–738. [\[CrossRef\]](#)
10. Beniwal, A.; Dadhich, R.; Alankar, A. Deep learning based predictive modeling for structure–property linkages. *Materialia* **2019**, *8*, 100435. [\[CrossRef\]](#)
11. Yang, Z.; Yabansu, Y.C.; Al-Bahrani, R.; Liao, W.K.; Choudhary, A.N.; Kalidindi, S.R.; Agrawal, A. Deep learning approaches for mining structure–property linkages in high contrast composites from simulation datasets. *Comput. Mater. Sci.* **2018**, *151*, 278–287. [\[CrossRef\]](#)
12. Liu, R.; Agrawal, A.; Liao, W.K.; Choudhary, A.; De Graef, M. Materials discovery: Understanding polycrystals from large-scale electron patterns. In Proceedings of the 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, USA, 5–8 December 2016; pp. 2261–2269.
13. Yang, Z.; Yabansu, Y.C.; Jha, D.; Liao, W.K.; Choudhary, A.N.; Kalidindi, S.R.; Agrawal, A. Establishing structure–property localization linkages for elastic deformation of three-dimensional high contrast composites using deep learning approaches. *Acta Mater.* **2019**, *166*, 335–345. [\[CrossRef\]](#)
14. Cecen, A.; Dai, H.; Yabansu, Y.C.; Kalidindi, S.R.; Song, L. Material structure–property linkages using three-dimensional convolutional neural networks. *Acta Mater.* **2018**, *146*, 76–84. [\[CrossRef\]](#)
15. Alqahtani, N.; Armstrong, R.T.; Mostaghimi, P. Deep learning convolutional neural networks to predict porous media properties. In Proceedings of the SPE Asia Pacific Oil and Gas Conference and Exhibition, Society of Petroleum Engineers, Brisbane, Australia, 23–25 October 2018.
16. Rong, Q.; Wei, H.; Huang, X.; Bao, H. Predicting the effective thermal conductivity of composites from cross sections images using deep learning methods. *Compos. Sci. Technol.* **2019**, *184*, 107861. [\[CrossRef\]](#)
17. Steinmetz, P.; Yabansu, Y.C.; Hötzer, J.; Jainta, M.; Nestler, B.; Kalidindi, S.R. Analytics for microstructure datasets produced by phase-field simulations. *Acta Mater.* **2016**, *103*, 192–203. [\[CrossRef\]](#)
18. Steinbach, I. Phase-Field Model for Microstructure Evolution at the Mesoscopic Scale. *Annu. Rev. Mater. Res.* **2013**, *43*, 89–107. [\[CrossRef\]](#)
19. Azimi, S.M.; Britz, D.; Engstler, M.; Fritz, M.; Mücklich, F. Advanced Steel Microstructural Classification by Deep Learning Methods. *Sci. Rep.* **2018**, *8*, 2128. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Liu, Y.; Zhao, T.; Ju, W.; Shi, S. Materials discovery and design using machine learning. *J. Mater.* **2017**, *3*, 159–177. [\[CrossRef\]](#)
21. Pilania, G.; Wang, C.; Jiang, X.; Rajasekaran, S.; Ramprasad, R. Accelerating materials property predictions using machine learning. *Sci. Rep.* **2013**, *3*, 2810. [\[CrossRef\]](#)
22. Meng, L.; McWilliams, B.; Jarosinski, W.; Park, H.Y.; Jung, Y.G.; Lee, J.; Zhang, J. Machine Learning in Additive Manufacturing: A Review. *JOM* **2020**, *72*, 2363–2377. [\[CrossRef\]](#)
23. Aggour, K.S.; Gupta, V.K.; Ruscitto, D.; Ajdelsztajn, L.; Bian, X.; Brosnan, K.H.; Chennimalai Kumar, N.; Dheeradhada, V.; Hanlon, T.; Iyer, N.; et al. Artificial intelligence/machine learning in manufacturing and inspection: A GE perspective. *MRS Bulletin* **2019**, *44*, 545–558. [\[CrossRef\]](#)
24. Pant, P.; Sjöström, S.; Simonsson, K.; Moverare, J.; Proper, S.; Hosseini, S.; Luzin, V.; Peng, R. A Simplified Layer-by-Layer Model for Prediction of Residual Stress Distribution in Additively Manufactured Parts. *Metals* **2021**, *11*, 861. [\[CrossRef\]](#)
25. Wang, C.; Tan, X.; Tor, S.; Lim, C. Machine learning in additive manufacturing: State-of-the-art and perspectives. *Addit. Manuf.* **2020**, *36*, 101538. [\[CrossRef\]](#)
26. Sakib, S.; Ahmed, N.; Kabir, A.J.; Ahmed, H. An Overview of Convolutional Neural Network: Its Architecture and Applications. 2019. Available online: <https://www.preprints.org/manuscript/201811.0546/v4> (accessed on 22 July 2021).
27. Traore, B.B.; Kamsu-Foguem, B.; Tangara, F. Deep convolution neural network for image recognition. *Ecol. Inform.* **2018**, *48*, 257–268. [\[CrossRef\]](#)
28. Al-Saffar, A.A.M.; Tao, H.; Talab, M.A. Review of deep convolution neural network in image classification. In Proceedings of the 2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), Jakarta, Indonesia, 23–24 October 2017; pp. 26–31.
29. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
30. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [\[CrossRef\]](#)
31. Ringnér, M. What is principal component analysis? *Nat. Biotechnol.* **2008**, *26*, 303–304. [\[CrossRef\]](#)
32. Abdi, H.; Williams, L.J. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 433–459. [\[CrossRef\]](#)

-
33. Bro, R.; Smilde, A.K. Principal component analysis. *Anal. Methods* **2014**, *6*, 2812–2831. [\[CrossRef\]](#)
 34. Li, Y.; Wu, H. A clustering method based on K-means algorithm. *Phys. Procedia* **2012**, *25*, 1104–1109. [\[CrossRef\]](#)
 35. Yuan, C.; Yang, H. Research on K-value selection method of K-means clustering algorithm. *J—Multidiscip. Sci. J.* **2019**, *2*, 226–235. [\[CrossRef\]](#)
 36. Quey, R. Neper Reference Manual. No. May **2019**. Available online: <https://neper.info/docs/neper.pdf> (accessed on 22 July 2021).
 37. Lecomte, J.S.; Bunge, H.J. *Texture Analysis in Materials Science H.-J. Bunge*; Elsevier: Amsterdam, The Netherlands, 2015. [\[CrossRef\]](#)
 38. Zhang, Y.; Chen, H.X.; Duan, L.; Fan, J.B.; Ni, L.; Ji, V. A comparison study of the structural and mechanical properties of cubic, tetragonal, monoclinic, and three orthorhombic phases of ZrO₂. *J. Alloy. Compd.* **2018**, *749*, 283–292. [\[CrossRef\]](#)
 39. Zamkovskaya, A.; Maksimova, E. Anisotropic thermal expansion in crystals of different categories. *J. Phys. Conf. Ser.* **2015**, *661*, 012067. [\[CrossRef\]](#)
 40. Permann, C.J.; Gaston, D.R.; Andrš, D.; Carlsen, R.W.; Kong, F.; Lindsay, A.D.; Miller, J.M.; Peterson, J.W.; Slaughter, A.E.; Stogner, R.H.; et al. MOOSE: Enabling massively parallel multiphysics simulation. *SoftwareX* **2020**, *11*, 100430. [\[CrossRef\]](#)
 41. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
 42. Nwankpa, C.E. Advances in optimisation algorithms and techniques for deep learning. *Adv. Sci. Technol. Eng. Syst. J.* **2020**, *5*, 563–577. [\[CrossRef\]](#)
 43. Grandini, M.; Bagli, E.; Visani, G. Metrics for Multi-Class Classification: An Overview. *arXiv* **2020**, arXiv:2008.05756.
 44. Novaković, J.D.; Veljović, A.; Ilić, S.S.; Papić, Ž.; Milica, T. Evaluation of classification models in machine learning. *Theory Appl. Math. Comput. Sci.* **2017**, *7*, 39–46.
 45. Tammina, S. Transfer learning using VGG-16 with deep convolutional neural network for classifying images. *Int. J. Sci. Res. Publ.* **2019**, *9*, 143–150. [\[CrossRef\]](#)
 46. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
 47. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
 48. Jayech, K. Regularized deep convolutional neural networks for feature extraction and classification. In *International Conference on Neural Information Processing*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 431–439.
 49. Siddique, R.; Aggarwal, P.; Aggarwal, Y. Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks. *Adv. Eng. Softw.* **2011**, *42*, 780–786. [\[CrossRef\]](#)
 50. Revi, V.; Kasodariya, S.; Talapatra, A.; Pilania, G.; Alankar, A. Machine Learning Elastic Constants of Multi-Component Alloys. *Comput. Mater. Sci.* **2021**, *198*, 110671. [\[CrossRef\]](#)