Available online at www.sciencedirect.com

Manufacturing Letters

Manufacturing Letters 33 (2022) 808–816

50th SME North American Manufacturing Research Conference (NAMRC 50, 2022)

Process Monitoring: Distinguishing Defect Shapes by Strain Field Signatures

Saurabh Basu[1,*], Christopher C. McComb[2], Mustafa Rifat[1], Hongtao Sun[1], Soundar Kumara[1]

[1]*Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, PA, USA*

[2]*Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA*

* Corresponding author. Tel.: +1 814 863 2447; fax: +1 814 863 4745. *E-mail address:* sxb514@psu.edu

Abstract

In this research, a theoretic physics-based framework for identification of defects via analysis of strain fields is presented. This framework comprises identification of self-similarity of strain fields followed by their dimensionality reduction using kernel based principal component analysis. The efficacy of this framework is tested qualitatively, by visual analysis, and quantitatively, using numerical classification algorithms. We see high (>95%) accuracy of classification via cross-validation studies using support vector machine algorithm. These results suggest that strain field can provide a viable approach for constructing highly robust in line defect detection system in modern manufacturing environments.

Keywords: Monitoring; Defect; Strain fields

1. Introduction

Functional requirements of modern appliances can be achieved using a combination of advanced geometries, and microstructures. The fabrication of such devices is often complicated by instabilities in the response of materials to thermomechanical conditions of manufacturing. These instabilities produce volumetric defects that compromise functionality. This can be addressed using in-line monitoring of material behavior in pursuit of identifying the onset of defect generation followed by timely mitigation. However, this is a challenging endeavor: (i) manufacturing defects are occluded which makes it challenging to 'observe' their generation rapidly (e.g., optically) for timely mitigation, and (ii) they affect the behavior of the parent material in complex ways that are often not well understood. This makes it challenging to systematically formulate 'signatures' that may be used to detect defects [1, 2]. Traditional approaches for in-line monitoring involve detection of behavior anomalies instigated by defects in low-dimensional signals such as force, and acceleration. This approach is however susceptible to noise in manufacturing environments. Theoretically, these shortcomings can be alleviated if the physics of the material being manufactured are accounted for within the in-line monitoring framework itself.

In this work, we provide numerical proof-of-concept of a framework for in-line monitoring of material health that is directly based on the physics of its behavior. Specifically, we describe how the linear effects of volumetric defects can be delineated in pursuit of their classification. For demonstrating the proof-of-concept, we ask the fundamental question: can we use elastic strain fields as indicators of the morphology of underlying defects. Our proposed solution to this problem is based on self-similarity analysis of elastic deformation fields with the underlying hypothesis- the 2-point correlation function of an elastic deformation field contains information to classify anomalous features of the parent microstructure. We look at the morphology of defects because their characteristics: (i) can be correlated to their thermomechanical origins [4]. Thus, their detection can advise revision of process parameters, and (ii) have been previously shown to affect response [5]. Our approach to this solution is inspired from other works on statistical analysis of microstructures [6] and involves dimensionality reduction of the 2-point correlation functions of elastic deformation fields with the scientific hypothesis- that a

small number of principal directions in the domain of elastic deformation fields can classify a larger number of microstructure types.

In the present work, we limit ourselves to only numerical data that is generated using a linear finite element framework. This choice enables us to exercise precise control as we establish the proof of concept of our proposed methodology for characterization of defect types using strain field signatures. We however note that numerically generated data is bereft of noise which is ubiquitous in empirical data. We attempt to delineate some plausible compromising effects of noise on the efficacy of this methodology in the forthcoming sections.

Section 2 presents a review of relevant literature. Section 3 presents the short description of the numerical methods used in this study. Section 4 provides results, that are then discussed in section 5. Section 6 provides some immediate future directions of research before section 7 that summarizes this research and provides concluding remarks.

2. Literature Review

In-line monitoring systems comprise in-situ modules for acquiring data, modules for processing acquired data, modules for decision making based on the processed data, and finally, modules for feedback control of the manufacturing platform. Such systems must provide within a reasonable amount of time the likelihood that a manufacturing step resulted in a defect, and if possible, the nature of the defect. This information can then be used to pause the manufacturing process until the source of the defects is identified and mitigated. All sources of latency must be addressed in in-line monitoring systems to ensure high throughput. Herein, the bottlenecks in such systems often involve data acquisition, and processing. To address these issues, several studies have looked at efficient methodologies for data acquisition, and processing. These are summarized in the following two subsections.

2.1 Methodologies for data acquisition

Methodologies to acquire manufacturing data are advised by a combination of the platform configuration, process physics, and material state variables such as temperature, and density. These methodologies can be dominantly classified into two types, viz. low-energy and high-energy. A classic low energy source used for imaging defects involves ultrasonic imaging, wherein an ultrasonic sound wave is transmitted through the component volume. Interaction of the ultrasonic wave with internal defects such as cracks, and porosities results in the wave being reflected. Characteristics of the reflected wave can be recorded, and subsequently processed to infer the presence of the corresponding occluded defect features. Ultrasonic waves are often used for detecting defects in castings [7,8], and welding [9]. More recently, their use has been explored in detecting defect features in additively manufactured components [10,11].

A second low-energy defect characterization methodology involves optical imaging. Unlike ultrasonic waves that traverse through the volume of the component, optical imaging can only acquire information from its surface. Nonetheless, features captured within these images have been used to infer presence of volumetric microstructural defects, e.g., in welding [12], and rolling operations. More recently, several applications of optical defect detection have been realized in additive manufacturing (AM) for detecting volumetric defects [13,14]. This is possible because AM involves a layer-by-layer build routine. This enables the optical imaging of each individual layer wherein subsequent data-processing can indicate the genesis of volumetric defects on the present layer allowing for timely intervention. Infra-red waves have also been used to detect presence of volumetric defects such as internal voids [15,16]. Here, the premise underlying detection of volumetric defects such as voids is their ability to perturb the dynamics of heat, which then results in anomalous signatures in the infra-red field emanated by the component.

High energy sources for in-line monitoring involve electromagnetic penetration with X-Rays, e.g., via computed tomography [17-19], and neutron beams, e.g., via cyclotron beam sources [20]. These methodologies for in-line defect detection are however resource intensive which limits their use for investigations on fundamental mechanisms of material behavior during manufacturing.

2.2 Methodologies for processing acquired data

Crucial towards the implementation of an in-line defect monitoring system is rapid processing of acquired data to infer presence of defects in the component being fabricated. This technical challenge is often approached as detection of anomalous features in the processed data, whereby presence of a defect is inferred. To this end, considerable research has been performed with the objective- to classify a feature in the acquired data as anomalous, rapidly, and reliably.

Common implementations of this concept include analysis of spectrums of machining force signatures [21]. Here, the onset of defects, e.g., due to tool wear, often produces vibrations that occupy a distinct part of the spectrum that is classified as anomalous. This concept is often also used to identify changes in cutting mechanisms even with fresh tools with the objective- to classify the material itself. Other data-processing approaches for detecting defects include machine learning algorithms such as Support Vector Machine (SVM) [22-24], k-nearest neighbor clustering [25-27], various types of neural networks [28-33], and random forests [26, 34-34]. SVMs tend to work well when there is a clear separation between different data classes (e.g., samples with and without defects) and contend well with high dimensionality; however, the training time often increases substantially when there are a large number of samples. In contrast, the k-nearest neighbors algorithm simply averages over the k points nearest to a sample in order to render a classification, and as such has negligible training time but can be slow during inference. Neural networks of various types have also been useful in this domain, and are capable of representing extreme nonlinearity in the training data. Finally, random forests are a common ensemble technique, rendering a classification by combining the

individual classifications of a set of decision trees trained on different subsets of the data.

These data-processing methodologies are primarily used for analysing the raw acquired data in pursuit of classifying features as anomalous. Nonetheless, another intermediate processing step may be added before such defect detection algorithms are employed. An example of the same includes detection of defects in large building structures [36]. Here, time-separated images of the same structure are used to calculate strain fields, which characterize deformation to the building structure. From these fields, evolution of anomalous features in the building structure are subsequently deduced to characteristics its structural health.

Characterization of strain fields from images involves the use of image correlation algorithms for calculation of displacement fields, which are subsequently differentiated to produce strain fields [37,38]. Here, sources of inaccuracy are dominantly two-fold comprising: (i) resolution of the optics used for acquiring images of the structure, and (ii) noise associated with image acquisition. To this end, the accuracy of defect detection using hand-held optical devices, e.g., smartphones cameras has been studied [39]. It was shown that a Samsung Galaxy S4 phone that was released in year 2013 can detect certain kinds of defects that are larger than 20 μm in size. Since then, optical assemblies that are meant to augment the resolution of smartphone cameras have been created and are commercially available. Presumably, the resolution of smartphone cameras that are equipped with such optical assemblies is amplified, making them capable of detecting finer defects. This is used as a practical premise in the current study, that handheld devices can be used to detect defects fine defects whose length scales are close to the optical limit.

3. Methods

3.1 Generating of numerical strain fields

The data used for this research comprised numerically generated strain fields of ASTM-E8 specimens that were endowed with volumetric defects and subsequently subject to elastic tensile deformation in numerical computing software Abaqus. This data was generated using the Python scripting interface of Abaqus and is available in an opensource repository [3]. The generation procedure involved creation of numerical ASTM-E8 plane strain tensile specimens featuring dimensions 6 mm × 38.1 mm. These tensile specimens were embedded with circular, elliptical, rectangular, triangular, peanut, and crescent shaped defects to produce respective defect area fractions of 7% − 9%. These numerical specimens were subsequently subject to tensile deformation of ϵ∼0.003 in the plane strain mode. Subsequently, the strain field components along the tensile axis were extracted at nodal points and discretized onto a square grid featuring a resolution of 0.1 mm /$pixel$. Zones that lie within the defects were assigned a strain $\epsilon = 0$. A total of 300 data specimens were generated for the present research (50 per defect type).

Fig. 1 shows instances of the strain field ϵ_{11} obtained from plane strain tensile deformation of the numerical ASTM E8 specimens featuring various defect types. These specimens were deformed to the same global strain of $\epsilon \sim 0.003$. However, it is qualitatively evident from these fields that the nature of the underlying defect affects it. This effect is most obvious via comparison of the crescent defect specimen with the triangular defect specimen. The former features a concave shape with two sharp points where strain concentrations can be seen. The latter features a convex shape with three sharp corners. However, this triangular defect shape, along with its interaction with neighboring defects tends to produce strain concentrations at its base, rather than at all three corners as can be expected. These features are shown using white circles in Fig. 1.

3.2 Extracting 2-point correlations

The elastic strain fields produced using finite elements were subsequently processed to produce 2-point correlation functions in the frequency domain. The 2-point correlation of a function f is defined as [40]:

$$C = \langle f(\tilde{r})\, f(\tilde{r'}) \rangle = \frac{\int_\Omega f(\tilde{r}) f(\tilde{r}')d\tilde{r}}{A} \quad (1)$$

Here, $\tilde{r}$, and $\tilde{r'}$ are two vectors that lie in the domain of function f, $\langle\rangle$ refers to the spatial average that is expanded on the RHS of the equation as the integral over the domain Ω of the function f, divided by the size A of the domain. The 2-point correlation function C characterizes an averaged metric of how self-similar two different points in the domain of the function f are. The utility of the 2-point correlation has been explored in reconstruction of synthetic microstructures that have equivalent self-similarity as an empirical characterized sample [41]. These synthetically produced microstructures visually appear like the original empirical samples from which they were generated, and hence presumably also behave like the parent empirically characterized counterparts.

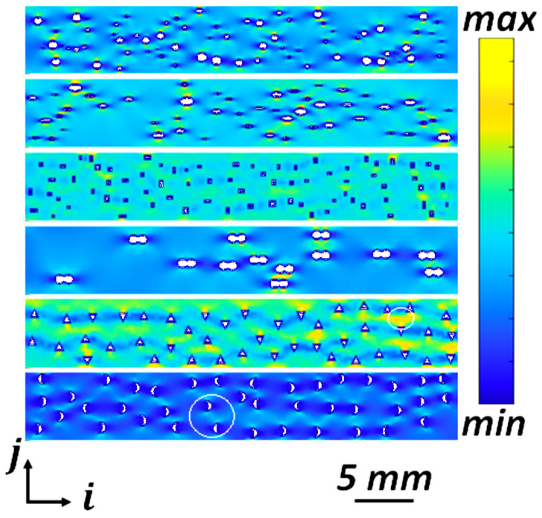

Figure 1. Strain fields ϵ_{11} obtained from simulated plane strain deformation of specimens featuring circular, elliptical, square, peanut, triangular, and crescent shape defects (top to bottom). The min of all strain fields is ϵ_{11}= 0, and max of the respective strain fields is ϵ_{11}= 5.8 ×10^{-3}, 5.8 ×10^{-3} ,4.7 ×10^{-3}, 3.8 ×10^{-3}, 6.8 ×10^{-3}, 8 ×10^{-3}, respectively.

In this study, it is hypothesized that 2-point correlations of strain fields implicitly contain signatures of the microstructures from which they originated. If this is true, the nature of the microstructure can be inferred by analysis of its strain fields, which is less resource intensive than direct characterization of the microstructures themselves. The former, e.g., characterization of strain fields can be achieved using optical characterization, in comparison with latter, e.g., characterization of microstructures, that often require sample preparation by polishing, and subsequent etching or electron microscopy. To this end, we also realize using the convolution theorem that the 2-point correlation of a function can be represented in the frequency domain concisely. Hence subsequent calculations are performed in the frequency domain. The convolution theorem is given by:

$$F(C) = F(f).\bar{F}(f) \tag{2}$$

where F represents the Fourier transform, $\bar{F}$ being its conjugate, and the RHS represents element wise multiplication of amplitudes of Fourier wave components of function f.

For software implementation of this concept, the ii component of the strain tensor is chosen arbitrarily. The reader is referred to the i direction in the reference axis in Fig. 1. This eventually enabled characterization of 2-point correlation functions using the eqn. 2 with $f = \epsilon_{ii}$. Results of this exercise are shown in Fig. 2.

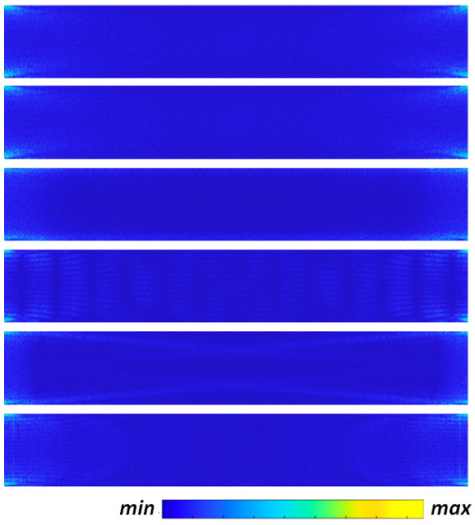

Figure 2. Autocorrelation functions for corresponding strain fields shown in Fig. 1, for circular, elliptical, square, peanut, triangular, and crescent shape defects (left to right). The color bar is logarithmic in scale. The (ln(1+min), ln(1+ max)) values of the respective autocorrelation are (−15.7,9.7), (−13.0, 9.4), (−15.4, 9.9), (−13.7, 9.9), (−14.9, 9.9), (−20.7, 9.6).

3.3. Classifying 2-point correlations using principal component analysis

Classification of the 2-point correlations of elastic strain fields was performed using kernel based principal component analysis [46,47]. In this approach, the original data points, e.g., frequency space representations of 2-point correlations of strain fields, e.g., $F(C)$ in eqn. 2, are mapped to a feature space $\Phi(C)$. This mapping enables the construction of the kernel matrix as a function of the dot product between the zero normalized projections of the αth, and βth data point:

$$K_{\alpha\beta} = \left(\left(\Phi(F(C_\alpha)) - \frac{1}{N}\sum_{a=1}^{N} \Phi(F(C_a))\right)^T \left(\Phi\left(F(C_\beta)\right) - \frac{1}{N}\sum_{a=1}^{N} \Phi(F(C_a))\right)\right)^\gamma \tag{3}$$

Here, $F(C_\alpha)$ is the frequency space representation of the 2-point correlation of the αth ϵ_{ii} strain field, in the columnar form, and T refers to transpose. For the present research, it was seen that a simple dot product operation with parameter $\gamma = 3$ was able to adequately classify the strain fields. Herein, $\Phi(F(C_\alpha))$ is the columnized form of the feature representation of data point, e.g., $\Phi(F(C_\alpha)) = F(C_\alpha)$ in this work. Finally, N is the number of samples used for constructing the kernel matrix. Of the total number of data points generated, e.g., 300, $N = 180$ samples (e.g., 30 for each microstructure type) were used for constructing the kernel matrix. This implies that dimensionality of the kernel matrix itself is 180×180.

Construction of the kernel matrix enables classification of the frequency space representations of the 2-point correlations of the test strain fields $F(C_{test})$ (e.g., 20 for each microstructure type). This involves finding the eigen values of the kernel matrix. Subsequently, the projections of the feature space mappings of the 'test' 2-point correlations $\Phi(F(C_{test}))$ are obtained with the eigen vectors. For the same, $\Phi(F(C_{test}))$ is represented as a linear sum of its projections on the respective basis vectors $\Phi(F(C_\alpha))$ that were used to construct the kernel matrix. Here, $\alpha \in \{1,2,..180\}$ wherein a valid projection of the feature space representations of the test data, viz. $\Phi(F(C_{test}))$ can be obtained on the Eigen vectors. To obtain these projections, the two eigen vectors with the largest two eigen values were chosen.

4. Results

4.1 Eigen values of the kernel matrix

The nature of these principal components obtained from the kernel matrix was evaluated by finding their cumulative fractional eigen values, or the total amount of variance that is captured from the original data using a progressively larger number of principal components. Fig. 3 shows the results of this analysis. It is seen that even the first principal component can account for > 99% of the variance in the feature space projections of the strain fields. This is expected given that defects account for only 7%-9% of the area fraction of the specimens, wherein the strain fields comprise fundamentally similar characteristics with subtle differences that are captured by a small number of additional principal components. Further, the variance captured by subsequent principal components decrease monotonically.

4.2 Classification of microstructures

Towards understanding the efficacy of these principal components (PCs) at classification of microstructures, the projections of the 2-point correlation functions $F(C_a)$ were plotted with respect to the first two eigen vectors of the kernel matrix. These vectors are hereafter referred as PC1, and PC2, respectively. Results are shown in Fig. 4, where hollow/solid markers refer to training/testing datasets, respectively. It is seen that the first principal component PC1 can sufficiently distinguish between several defect shapes. Strain fields originating from specimens containing crescent defects cluster towards the left at smaller values of PC1 (green diamonds). In comparison, strain fields originating from square, peanut, and elliptical shapes (squares, stars, and inverted triangles, respectively) cluster towards the right, e.g., at larger values of PC1 and show greater spreads along PC2. Interestingly, peanut, ellipse, and rectangular data points were seen to feature similar values of PC1. Origins of this behavior can be expected to lie in the similarity across these defect shapes in as it pertains to their mechanical behaviour.

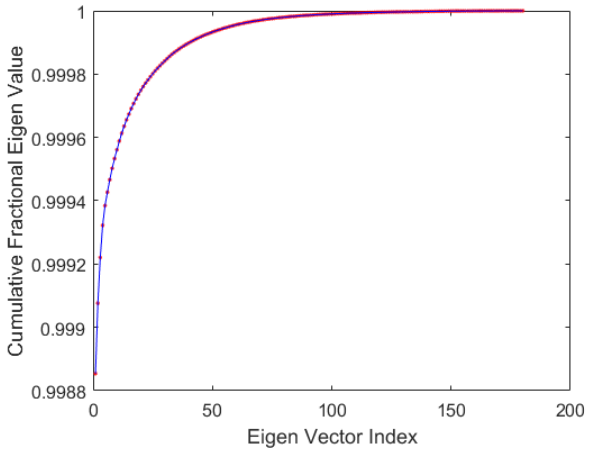

Figure 3. Fractional variance accounted for by principal components of 2-point correlation fields (cf. Fig. 2.).

Fig. 4 visually confirms that strain fields contain sufficient details for classification of numerically generated specimens based on the shapes of their defects. The robustness of this concept for identification of microstructure defect features were quantified by training a series of Error Correcting Output Codes (ECOC) Support Vector Machine (SVM) model [42,48], using 10-fold cross-validation [49] on the data identified in earlier sections. It was seen that not only are the clustering patterns in Fig. 4, qualitatively clear, but they also have quantitative value. When trained to classify the defect shapes using 2 principal components, 92.5±7.3% (mean ± standard deviation) of specimens in the validation fold are correctly classified. In order to construct a confusion matrix that shows how results are misclassified, we retrain a single ECOC SVM based with a 50% train/test split, and compute the confusion matrix based on the testing data. These results are displayed in a confusion matrix (see Fig. 5). We can observe that the model misclassifies circular defects as triangular, and also peanut-shaped defects as elliptical. These categories are adjacent to one another in Figure 4, which provides some face validity for the confusion matrix.

5. Discussion

Results in section 2 suggest that elastic strain fields expressed by a parent specimen when subject to external loads are distorted in the presence of defects. These distortions manifest signatures in the frequency space of the strain field. In this section we explore an explanation for this characteristic. This explanation is inspired from Mura's eigen-strain formulation for finding residual stresses σ_{ij} because of eigen-strain inhomogeneities ϵ^*_{ij} dispersed in the matrix [43]. In this formulation, the total strain that results from these inhomogeneities is linearly decomposed as a summation of the eigen-strains themselves ϵ^*_{ij}, and the resulting elastic strains e_{ij}. This is given by equation:

$$\epsilon_{ij} = e_{ij} + \epsilon^*_{ij} \tag{4}$$

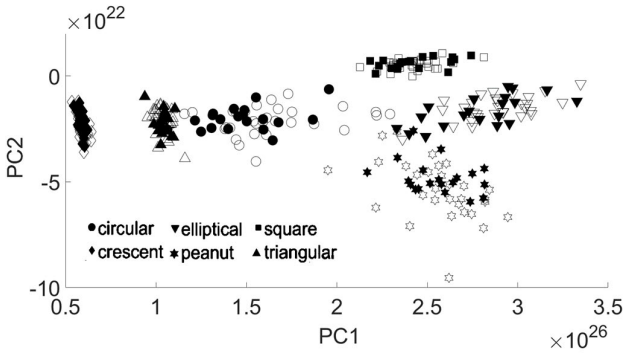

Figure 4. Projections of 2-point correlation windows on principal components PC1 and PC2. The hollow markers refer to the 180 training datasets (30 X 6), whereas the solid markers refer to 120 testing datasets (20 X 6).

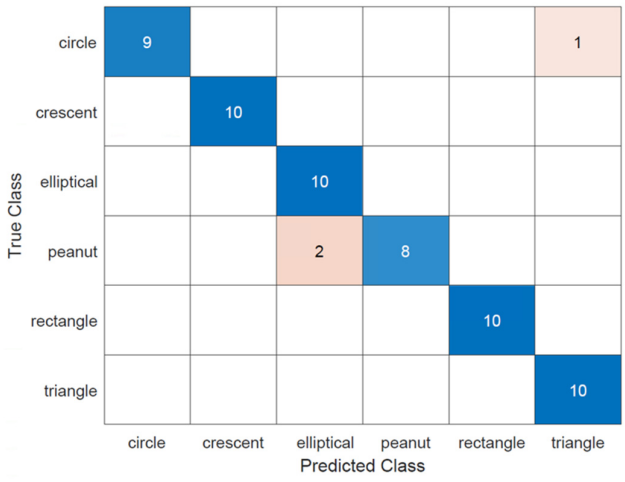

Figure 5. Confusion Matrix created for ECOC SVM results obtained from kernel PCA shown in Fig. 4.

The methodology for solving for e_{ij} in eqn. 4 relies on calculation of force equilibrium, i.e.,

$$\sigma_{ij,j} = 0 \; \forall \; i \tag{5}$$

The stress field σ_{ij} is provided using the elastic strain field e_{kl}, and the elastic modulus C_{ijkl} as $\sigma_{ij} = C_{ijkl}e_{kl}$. This is substituted in eqn. 5, resulting in:

$$C_{ijkl}e_{kl,j} = 0 \tag{6}$$

Using eqn. 4, we get:

$$e_{kl} = \epsilon_{kl} - \epsilon^*_{kl} \tag{7}$$

and

$$C_{ijkl}e_{kl} = C_{ijkl}(\epsilon_{kl,j} - \epsilon^*_{kl,j}) = 0 \tag{8}$$

A non-trivial solution to this equation is found by linearly decomposing the eigen-strain components in terms of their Fourier wave amplitudes, i.e.,

$$\epsilon^*_{kl} = \Sigma \bar{\epsilon}_{kl}{}^*(\tilde{\xi}) \exp(i\tilde{\xi} \cdot \tilde{x})$$

$$\Rightarrow \epsilon^*_{kl,j} = i\xi_j \Sigma \bar{\epsilon}^*(\tilde{\xi}) \exp(i\tilde{\xi} \cdot \tilde{x})$$

Here, $i = \sqrt{-1}$, and $\tilde{\xi}.\tilde{x} = \xi_k x_k$. Expressing total strain as a derivative of displacement, e.g., $\epsilon_{kl} = \partial u_k / \partial x_l$ enables us to solve for total displacement as a function of its corresponding Fourier components as:

$$\epsilon_{kl} = \partial \Sigma \bar{u}_k(\tilde{\xi}) \exp(i\tilde{\xi} \cdot \tilde{x}) / \partial x_l \; = i\xi_l \, \Sigma \bar{u}_k \exp(i\tilde{\xi} \cdot \tilde{x}) \tag{9}$$

Thus, $\epsilon_{kl,j} = \frac{\partial^2 u_k}{\partial x_l \partial x_j} = -\xi_l \xi_j \, \Sigma \bar{u}_k(\tilde{\xi}) \exp(i\tilde{\xi} \cdot \tilde{x})$, which after substituting in equation 8 provides:

$$-C_{ijkl}\xi_l \xi_j \bar{u}_k(\tilde{\xi}) \exp(i\tilde{\xi} \cdot \tilde{x}) \; = iC_{ijkl}\xi_j \bar{\epsilon}^*_{kl}(\tilde{\xi}) \exp(i\tilde{\xi} \cdot \tilde{x}) \tag{10}$$

The solutions to $\bar{u}_k$ are obtained from this equation. From these wave amplitudes, the displacement field components u_k are reconstructed as a Fourier series as: $u_k = \Sigma \bar{u}_k \exp(i\tilde{\xi} \cdot \tilde{x})$. Subsequently, strain field components $\epsilon_{kl} = \partial u_k / \partial x_l$ can be constructed by differentiating these displacement components. It is noted here that this argument, in the absence of surface traction equilibrium, viz., $n_i \sigma_{ij} = 0 \; \forall \; j$, conforms only to an infinite medium unlike our finite sized numerical ASTM E8 specimens. Nonetheless, the sizes of the defects are small compared with the size of the specimen, wherein the validity of this argument for finite sized samples may be implicitly assumed. This surface traction condition is of course naturally implemented in the finite element framework from which the strain fields shown in Fig. 1 were produced.

This argument suggests that the strain fields originating from the interaction of voids with the globally imposed stress may be modelled as consequences of fictitiously imposed eigen strains. The use of fictitious eigen-strains to analyze total strain fields is well known [44,45]. These eigen strains must have sufficiently unique characteristics such that the resulting total strain fields ϵ_{kl}, can be clustered (cf. Fig. 4). Further, the argument also shows that these strain fields may be completely represented in terms of wave amplitudes, e.g., as a spectrum. Herein, we note that from Eqn. 2 that 2-point correlations are simply squares of the amplitudes of the corresponding wave components in the spectrum. In this regard, the process of identifying principal components and clustering strain fields is equivalent to clustering the respective fictitious eigen-strains that implicitly contain characteristics specific to the geometry of the defects.

Some consequences of these insights are demonstrated here. The first involves a study of the effect of artificial changes to the spectrum of a strain field on its ability to be clustered accurately. Such a change can be produced by blurring of the strain field. This can be achieved using a low pass filter of the form $a' = ae^{-\alpha d(\tilde{\xi})}$, applied in the frequency space. Here a refers to the original amplitude of the 2-dimensional wave component of frequency $\tilde{\xi}$, a' being the filtered amplitude, $d(\tilde{\xi})$ the Euclidean distance of the frequency component $\tilde{\xi}$ from the origin of the frequency space, and $\alpha \in \{0.05, 0.2\}$ is an arbitrary parameter. Note, a larger α parameter emulates a greater level of blurring. Figs. 6 and 7 show the consequences of this filtering using parameters α = 0.05 and α = 0.2, respectively. These images were obtained by inverse Fourier transformation of the filtered frequency space. For facilitating comparison, the strain fields in these figures correspond to those in Fig. 1, respectively.

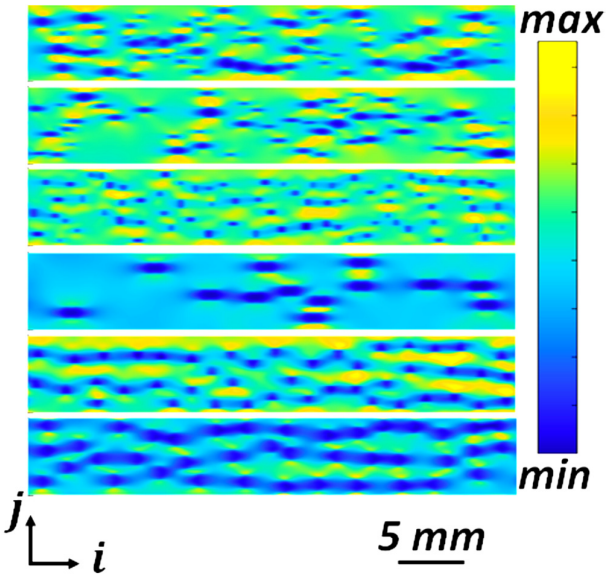

Figure 6. Strain fields ϵ_{11} obtained from filtering (α = 0.05) of simulated plane strain deformation of specimens featuring circular, elliptical, square, peanut, triangular, and crescent shape defects (top to bottom).

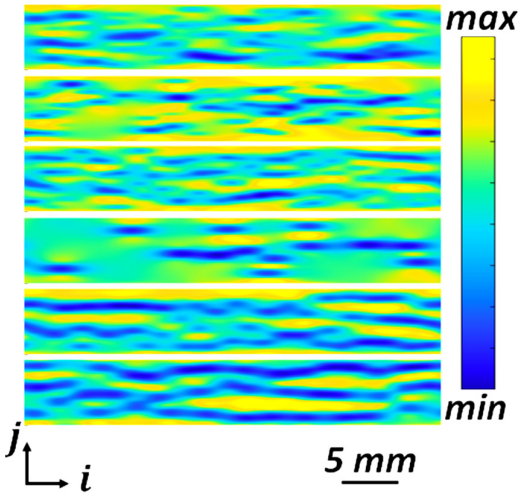

Figure 7. Strain fields ϵ_{11} obtained from filtering ($\alpha = 0.2$) of simulated plane strain deformation of specimens featuring circular, elliptical, square, peanut, triangular, and crescent shape defects (top to bottom).

As expected, a larger value of parameter α results in greater blurring. It is interesting that compared with the filter featuring $\alpha = 0.05$ (e.g., Fig. 5), the filter with $\alpha = 0.2$ (e.g., Fig. 6) blurs the strain fields beyond recognition through the naked eye. Herein, the shapes of the defects that are still recognizable in Fig. 5 are barely (if at all) recognizable in Fig. 6. The filtered strain fields were also analyzed using the approach described in section 3.3. A new set of principal components was generated for each case. Fig. 8 and 9 show the results of this analysis for filter parameter α = 0.05, and 0.2, respectively. This analysis suggests that blurring of the image causes only subtle changes to the ability of strain fields to be clustered. This inference is made simply by visual inspection, e.g., the clusters corresponding to different microstructure types are still uniquely visible. This implies that the dominant aspects of strain fields are accommodated by wave components featuring larger wavelengths, e.g., in their Fourier space representations. We do note that for filtering with $\alpha = 0.2$ (cf. Fig. 9), the peanut, and elliptical shapes show considerably larger overlap along the PC2 direction than the case with $\alpha = 0.05$ (cf. Fig. 8), or $\alpha = 0$ (cf. Fig. 4). This observation generally highlights the importance of the 'cleanliness' of the acquired data for classification of microstructures.

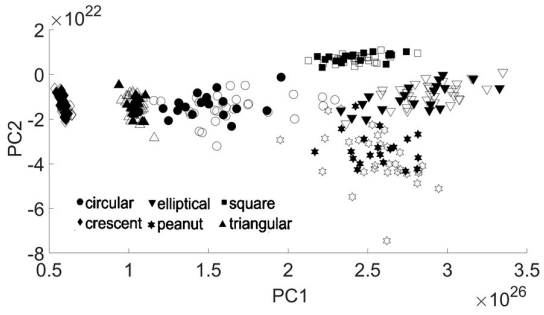

Figure 8. Projections of 2-point correlations $F(C)$ obtained from their filtered strain fields ($\alpha = 0.05$) on principal components PC1 and PC2. The hollow/solid markers refer to training/testing datasets.

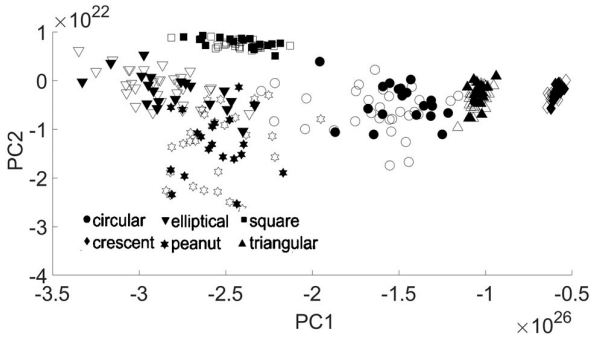

Figure 9. Projections of 2-point correlations $F(C)$ obtained from their filtered strain fields ($\alpha = 0.2$) on principal components PC1 and PC2. The hollow/solid markers refer to training/testing datasets.

The argument provided in this section so far also implies that the use of the raw strain field directly for clustering, in comparison with its 2-point correlation will result in sub-par classification efficacy. This is because the raw strain field is highly susceptible to phase of its wave components. This is an expected result. We note however that the 2-point correlation filters out phase information [41], making the proposed methodology insensitive to phase effects. To appreciate this thought, Fig. 10 shows the clusters obtained from the strain fields directly. A comparison of this figure with Fig. 4 suggests that the raw strain field plots cannot be used to accurately differentiating between square, triangular, and peanut shape defects. These clusters are seen to completely overlap in Fig. 10. This is also reflected in the quantified efficacy of clustering, e.g., the confusion matrix (see Fig. 11), that showed approximately 30%, 50%, 30%, and 40% error rates in classification of microstructures comprising circle defects, elliptical defects, peanut defects, and square defects, respectively.

Interestingly, the use of the 2-point correlation of the defect structure for their classification also produces poor results, compared with the 2-point correlation of strain fields. The corresponding confusion matrix is shown in Fig. 12. Here, errors of 40%, 40%, and 100% are seen in classification of circle, ellipse, and square/rectangular defects, respectively, without any filtering (α=0). This observation also validates our methodology for characterization of microstructures using strain fields which is based on the hypothesis that strain fields contain signatures that are highly correlated to characteristics of the defect which lie in the matrix.

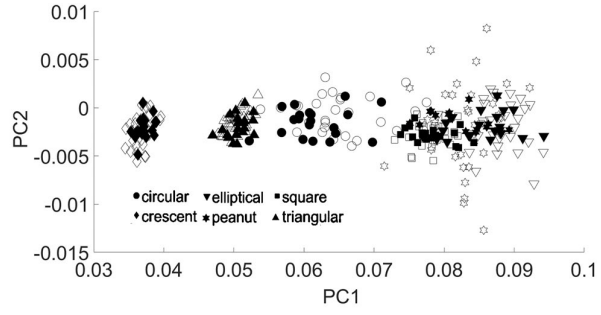

Figure 10. Direct projections of 2-point correlations of defect fields on principal components obtained using kernel PCA (cf. section 3.3). The hollow/solid markers refer to training/testing datasets.

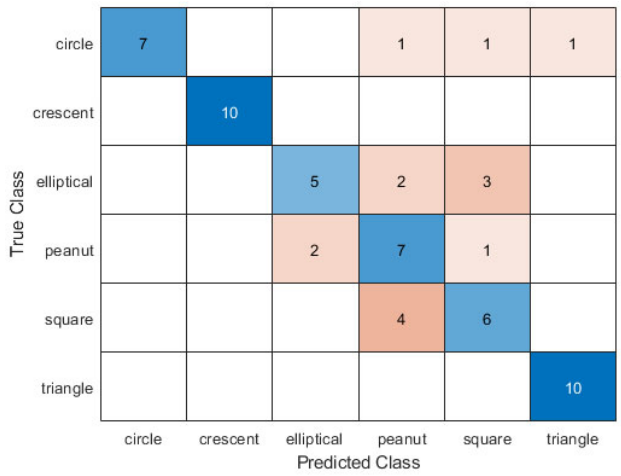

Figure 11. Confusion Matrix created for ECOC SVM results obtained from kernel PCA shown in Fig. 10.

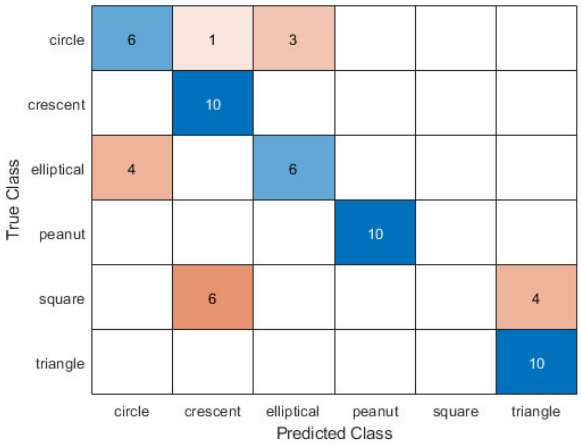

Figure 12. Confusion Matrix created for ECOC SVM results obtained from kernel PCA of 2-point correlation of defect structures directly.

6. Future Directions

The work described in this manuscript looks at the use of strain fields for classifying their underlying microstructures. It is hypothesized that the proposed methodology can be used synergistically with optical characterization of strain fields. To this end, this work naturally warrants answers to several key questions that the authors are currently in the process of addressing.

i. The proposed methodology assumes strain fields are acquired from a zone that is rich with defects. In this regard, the efficacy of this approach in zones that only comprise sparse, or even non-homogenous densities of these defects is not known. It is also not known how rapidly the efficacy of this approach will degrade if the defects are smaller in size, which would progressively make it more challenging to characterize strain fields accurately.

ii. The proposed methodology also assumes that defects are directly visible. The presence of defects in the close vicinity beneath a surface can affect the strain fields that form directly on the exposed surface of the component during external loading. In this regard, it is expected that this approach can classify even those defects that are hidden from the line of sight. However, the accuracy of this approach is unknown in such scenarios and can be expected to degrade rapidly for defects found at progressively larger depths beneath the surface. A parametric study is required to delineate such effects.

iii. Even a weakly non-linear behavior of the matrix will affect the strain fields that result from external loading. The resulting effects on efficacy of classification of microstructures is not well understood.

7. Summary and Conclusions

The present work delineates how effect of defects on strain fields during external loading of the parent component may be analysed to classify the underlying defect shapes. Towards this, elastic plane strain deformation of specimens embedded with 6 different defect types was simulated using finite element analysis. The resulting strain fields were processed for identification of their self-similarity using 2-point correlation functions. A reduced dimension representation of the 2-point correlation was then constructed using kernel based principal component analysis. The efficacy of this reduced dimension representation at classification of defect shapes was then tested visually, and quantitatively. It was realized that this theoretical approach was accurate for >95% of all specimens. This efficacy was degraded by:

i. Blurring of strain field, and thereby artificially changing the spectrum
ii. Direct use of strain field compared with the frequency space representation of its 2-point correlation
iii. Direct use of defect microstructure without any strain field information.

Plausible underlying reasons that enable the classification of defective microstructures using the frequency space representation of the 2-point correlations of their strain fields were discussed. Mura's eigen-strain formulation was used for the same. Finally, some future directions of this work were also described.

Acknowledgements

SB would like to acknowledge support of NSF grant #1825686. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF. SB and CCM acknowledge the support of the Institute of Computing and Data Sciences at PSU. SK acknowledges Pearce Professorship.

References

[1] Teti, R., Jemielniak, K., O'Donnell, G. and Dornfeld, D., 2010, Advanced monitoring of machining operations, Annals of the CIRP, 59(2):717-739.

[2] Suh, J.H., Kumara, S.R. and Mysore, S.P., 1999, Machinery fault diagnosis and prognosis: application of advanced signal processing techniques, Annals of the CIRP, 48(1):317-320.

[3] Khanolkar, P., Basu, S. and McComb, C., 2020, Image-based data on strain fields of microstructures with porosity defects, Data in Brief, 106627.

[4] Bourell, D., Kruth, J.P., Leu, M., Levy, G., Rosen, D., Beese, A.M. and Clare, A., 2017, Materials for additive manufacturing, Annals of the CIRP, 66(2):659-681.

[5] Van Hooreweder, B., Kruth, J.P., 2017, Advanced fatigue analysis of metal lattice structures produced by Selective Laser Melting, Annals of the CIRP, 66(1):221-224.

[6] Fullwood, D.T., Niezgoda, S.R., Adams, B.L. and Kalidindi, S.R., 2010, Microstructure sensitive design for performance optimization, Progress in Materials Science, 55(6):477-562.

[7] Hossain, M. S., Krenek, R., Taheri, H., & Dababneh, F. (2020, November). Ultrasonic Phased Array Technique for Defect Detection and Sizing in Heavy-Walled Cast Components. In ASME International Mechanical Engineering Congress and Exposition (Vol. 84492, p. V02BT02A038). American Society of Mechanical Engineers.

[8] Kawashima, K., Ito, T., & Nagata, Y. (2010). Detection and imaging of nonmetallic inclusions in continuously cast steel plates by higher harmonics. *Japanese Journal of Applied Physics*, *49*(7S), 07HC11.

[9] Stares, I. J., Duffil, C., Ogilvy, J. A., & Scruby, C. B. (1990). On-line weld pool monitoring and defect detection using ultrasonic. NDT international, 23(4), 195-200.

[10] Zhang, J., Wu, J., Zhao, X., Yuan, S., Ma, G., Li, J., ... & Ding, H. (2020). Laser ultrasonic imaging for defect detection on metal additive manufacturing components with rough surfaces. *Applied Optics*, *59*(33), 10380-10388.

[11] Honarvar, F., & Varvani-Farahani, A. (2020). A review of ultrasonic testing applications in additive manufacturing: Defect evaluation, material characterization, and process control. *Ultrasonics*, *108*, 106227.

[12]Senthil Kumar, G., Natarajan, U., & Ananthan, S. S. (2012). Vision inspection system for the identification and classification of defects in MIG welding joints. *The International Journal of Advanced Manufacturing Technology*, *61*(9), 923-933.

[13]Abdelrahman, M., Reutzel, E. W., Nassar, A. R., & Starr, T. L. (2017). Flaw detection in powder bed fusion using optical imaging. Additive Manufacturing, 15, 1-11.

[14]Al Mamun, A., Liu, C., Kan, C., & Tian, W. (2021). Real-time process authentication for additive manufacturing processes based on in-situ video analysis. *Procedia Manufacturing*, *53*, 697-704.

[15]Zhang, X., Saniie, J., & Heifetz, A. (2020). Detection of defects in additively manufactured stainless steel 316L with compact infrared camera and machine learning algorithms. JOM, 72(12), 4244-4253.

[16] Bartlett, J. L., Heim, F. M., Murty, Y. V., & Li, X. (2018). In situ defect detection in selective laser melting via full-field infrared thermography. *Additive Manufacturing*, *24*, 595-605.

[17]Bellens, S., Vandewalle, P., & Dewulf, W. (2021). Deep learning based porosity segmentation in X-ray CT measurements of polymer additive manufacturing parts. Procedia CIRP, 96, 336-341.

[18]Ziółkowski, G., Chlebus, E., Szymczyk, P., & Kurzac, J. (2014). Application of X-ray CT method for discontinuity and porosity detection in 316L stainless steel parts produced with SLM technology. Archives of civil and mechanical engineering, 14(4), 608-614.

[19]Thompson, A., Maskery, I., & Leach, R. K. (2016). X-ray computed tomography for additive manufacturing: a review. Measurement Science and Technology, 27(7), 072001.

[20]González, G. B., Cohen, J. B., Hwang, J. H., Mason, T. O., Hodges, J. P., & Jorgensen, J. D. (2001). Neutron diffraction study on the defect structure of indium–tin–oxide. Journal of Applied Physics, 89(5), 2550-2555.

[21]Chen, S., Lin, B., Han, X., & Liang, X. (2013). Automated inspection of engineering ceramic grinding surface damage based on image recognition. *The International Journal of Advanced Manufacturing Technology*, *66*(1), 431-443.

[22]Gobert, C., Reutzel, E. W., Petrich, J., Nassar, A. R., & Phoha, S. (2018). Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging. Additive Manufacturing, 21, 517-528.

[23]Ge, M., Du, R., Zhang, G., & Xu, Y. (2004). Fault diagnosis using support vector machine with an application in sheet metal stamping operations. Mechanical Systems and Signal Processing, 18(1), 143-159.

[24]Malarvel, M., & Singh, H. (2021). An autonomous technique for weld defects detection and classification using multi-class support vector machine in X-radiography image. Optik, 231, 166342.

[25]Lin, J., Yang, J., Huang, Y., & Lin, X. (2021). Defect identification of metal additive manufacturing parts based on laser-induced breakdown spectroscopy and machine learning. Applied Physics B, 127(12), 1-10.

[26]Li, R., Jin, M., & Paquit, V. C. (2021). Geometrical defect detection for additive manufacturing with machine learning models. *Materials & Design*, *206*, 109726.

[27]Kumar, P., & Jain, N. K. (2022). Surface roughness prediction in micro-plasma transferred arc metal additive manufacturing process using K-nearest neighbors algorithm. *The International Journal of Advanced Manufacturing Technology*, 1-13.

[28]Gaja, H., & Liou, F. (2018). Defect classification of laser metal deposition using logistic regression and artificial neural networks for pattern recognition. *The International Journal of Advanced Manufacturing Technology*, *94*(1), 315-326.

[29]Lalithakumari, S., Sheelarani, B., & Venkatraman, B. (2012). Artificial neural network based defect detection of welds in tofd technique. International Journal of Computer Applications, 41(20).

[30]Duran, O., Althoefer, K., & Seneviratne, L. D. (2007). Automated pipe defect detection and categorization using camera/laser-based profiler and artificial neural network. IEEE Transactions on Automation Science and Engineering, 4(1), 118-126.

[31]Lee, C., Seo, G., Kim, D. B., Kim, M., & Shin, J. H. (2021). Development of Defect Detection AI Model for Wire+ Arc Additive Manufacturing Using High Dynamic Range Images. *Applied Sciences*, *11*(16), 7541.

[32]Westphal, E., & Seitz, H. (2021). A machine learning method for defect detection and visualization in selective laser sintering based on convolutional neural networks. Additive Manufacturing, 41, 101965.

[33]Shang, L., Yang, Q., Wang, J., Li, S., & Lei, W. (2018, February). Detection of rail surface defects based on CNN image recognition and classification. In 2018 20th International Conference on Advanced Communication Technology (ICACT) (pp. 45-51). IEEE.

[34]García-Moreno, A. I., Alvarado-Orozco, J. M., Ibarra-Medina, J., & Martínez-Franco, E. (2020). Image-based porosity classification in Al-alloys by laser metal deposition using random forests. *The International Journal of Advanced Manufacturing Technology*, *110*(9), 2827-2845.

[35]Zhang, Z., Yang, Z., Ren, W., & Wen, G. (2019). Random forest-based real-time defect detection of Al alloy in robotic arc welding using optical spectrum. Journal of Manufacturing Processes, 42, 51-59.

[36]Liu, R. M., Babanajad, S. K., Taylor, T., & Ansari, F. (2015). Experimental study on structural defect detection by monitoring distributed dynamic strain. *Smart Materials and Structures*, *24*(11), 115038.

[37]Rifat, M., DeMeter, E. C., & Basu, S. (2020). Microstructure evolution during indentation of Inconel-718 created by additive manufacturing. *Materials Science and Engineering: A*, *781*, 139204.

[38]Kulkarni, A., Vidvans, A., Rifat, M., Bicknell, G., Gong, X., Manogharan, G., ... & Basu, S. (2018, June). Scalable Detection of Defects in Additively Manufactured PLA Components. In *International Manufacturing Science and Engineering Conference* (Vol. 51357, p. V001T01A042). American Society of Mechanical Engineers.

[39]Vidvans, A., & Basu, S. (2018). Smartphone based scalable reverse engineering by digital image correlation. *Optics and Lasers in Engineering*, *102*, 126-135.

[40]Beran, M., 1965, Statistical continuum theories, Transactions of the Society of Rheology, 9(1):339-355.

[41]Fullwood, D. T., Niezgoda, S. R., & Kalidindi, S. R. (2008). Microstructure reconstructions from 2-point statistics using phase-recovery algorithms. *Acta Materialia*, *56*(5), 942-948.

[42]Dietterich, T. G., & Bakiri, G. (1994). Solving multiclass learning problems via error-correcting output codes. Journal of artificial intelligence research, 2, 263-286.

[43]Mura, T. (2013). *Micromechanics of defects in solids*. Springer Science & Business Media.

[44]Korsunsky, A. M. (2005). The modelling of residual stresses due to surface peening using eigenstrain distributions. *The Journal of Strain Analysis for Engineering Design*, *40*(8), 817-824.

[45]Jun, T. S., Venter, A. M., & Korsunsky, A. M. (2011). Inverse eigenstrain analysis of the effect of non-uniform sample shape on the residual stress due to shot peening. *Experimental Mechanics*, *51*(2), 165-174.

[46]Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, *2*(11), 559-572.

[47]Schölkopf, B., Smola, A. and Müller, K.R., 1998. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5), 1299-1319.

[48]Fürnkranz, Johannes. 2002. Round Robin Classification. *Journal of Machine Learning Research*, 2, 721–747.

[49]Stone M. 1974. Cross-validatory choice and assessment of statistical predictions. *J. Royal Stat. Soc.*, 36(2), 111–147.