Cluster-Based Colormap of Nanoindentation Using Machine Learning

Sarah Ghaleb Alhamdani, Abdallah Alrshdan and Khalid Ahmed
College of Engineering, Alfaisal University, Riyadh, Saudi Arabia
salhamdani@alfaisal.edu, aalrshdan@alfaisal.edu, kahmed@alfaisal.edu

Abstract

Nanoindentations is an advanced method of measuring the mechanical properties of small volumes of materials using an instrumented indentation technique. Properties such as elastic modulus and hardness can be measured and explored by Nanoindentation. To support the material designers and specialists, this research is aiming to segment the captured data into regions with homogenous physical properties that can be used to optimize the process of mixing composite materials. During the Nanoindentation tests, and depending on the grid size, the resulting image/map of mechanical properties from the tested material's layer has a large dimensionality variety of values. Image processing technique and Machine Learning algorithms were utilized to cluster the physical property map and reduce its dimensionality. The sample data for this study is data of SLA 3D printed Nanocomposite of acrylic polymer (boron nitride), which is used as a case study to apply denoising and clustering techniques.

Keywords
Nanoindentation, Image Denoising, Segmentation, K-Means Clustering and Machine Learning.

1. Introduction
Instrumented indentation or shortly Nanoindentation technique is an effective and powerful tool to detect the local elastic properties and hardness of the material microstructure and nanostructure. The resulted values of mechanical properties from the tested material's layer have huge variety of numbers; elastic properties and hardness values, even though the material size is very small -at nano-scale-. Nanoindentation is like any data capturing process where the accuracy of the captured data is a concern. The accuracy of the captured data depends on many factors which are mainly related to the device errors, environmental factors, and the subject where the data is collected from. These factors cause noises data that disturb the homogeneity of the physical property of the material tested and create incorrect classification and identification of material properties. A complete understanding of all mechanical properties is crucial to understanding the performance of a material and therefore good quality control and assurance of material's design.

1.1 Objectives
In this research, the physical property acquired from the nanoindentation process is treated as an image. In images, the value of the acquired data is mainly the color of each point (pixel). The objective of this research is to segment the captured data into regions with homogenous physical properties, by denoising image and using machine learning technique. This will allow accurate characterization of the material.

2. Literature Review
Nanoindentation is the most commonly used technique to study the microstructural mechanical properties of cement-based materials (Hu and Li 2015). It has also been particularly useful in the characterization of polymer matrix materials, fibers, particles, and other constituent materials in polymer composites (Ronald 2014). To date, nanoindentation is the most sought-after technique used to explore and study the microstructural behavior of cementitious composites (Gautham and Saptarshi 2019). Composite materials are regularly affected by different mixing conditions and components ratios (Chmielewski et al. 2010) and (Hu 2015). Despite the undoubted contribution of Nanoindentation to the measurements of mechanical properties; care must be taken during experiments to properly understand the obtained results while testing certain types of materials. Grid nanoindentation is a two-dimensional- mapping tool used for examining the properties of constituent phases independently of each other. (Nohava et al. 2010).

Different analysis methods were undertaken in analyzing grid indentation data, such as Statistical analysis and Clustering analysis. Statistical analysis was applied to analyze mechanical properties of the Bakken; along the Mori-Tanaka scheme was carried out to homogenize the elastic properties of the samples and upscale the nanoindentation data to the macroscale (Kouqi et al. 2018). A straightforward application of instrumented indentation to statistically extract the in situ elastic properties of individual components and to image the connectivity among these phases in composites; is applied and tested on a titanium-titanium monoboride (Ti–TiB) of various volumetric proportions (Constantinides et al. 2006). Static nanoindentation and Maximum Likelihood Estimation (MLE) were applied for the nano/micromechanical properties investigation of alkali-activated fly ash (AAFA) (Zhiyu et al. 2020).

Due to the limitation of Statistical Analysis, Clustering analysis was introduced to deal with grid nanoindentation data of cement paste samples. Compared with the deconvolution method and Gaussian mixture model, clustering analysis is better to avoid the normal distribution assumption, more robust for fewer data samples, and preferable in computational efficiency and stability (Chen et al. 2021).

The cluster-based analysis is an unsupervised Machine Learning method where a clustering algorithm determine the groups of data points without human interference, which helps in detecting hidden anomaly and abnormal data (Syarif et al. 2012). Since it's hard to tell one clustering algorithm is better than others (Nagpal 2013), K-means clustering is a wide-use algorithm in nanoindentation (Chen et al. 2021), (Koumoulos et al. 2019) and (Vignesh 2019). Another study used Supervised Machine Learning with KNN and SVM classification models (Koumoulos et al. 2020). K-Means Clustering algorithm involves the division of data into distinct classes where each class of data has significant properties. It is a strategy used for cluster investigation in which various observations are classified into k clusters and each of the observation belongs to the cluster with the nearest mean (Arshleen 2018).

Nanoindentation is like any data capturing process where the accuracy of the captured data is a concern. The accuracy of the captured data depends on many factors which are mainly related to the device errors, environmental factors, and the subject where the data is collected from. The nanoindentation process inherits different errors that degrade the accuracy of the acquired data. Menčík (2012) has listed many errors associated with the nanoindentation data acquisition process. The list includes but is not limited to properties of the tested material, device properties, temperature changes, and humidity. These factors cause noises data that disturb the homogeneity of the physical property of the material tested. Noise cleaning is very essential to increase the accuracy of captured data. In image processing, denoising is very common to improve image quality. Filtering methods are the most common techniques to denoise 2D or 3D images. Goyal et al. (2020) classified and compared some of the significant techniques in the field of denoising.

3. Methods

The research methodology includes three main steps: data acquisition, denoising, and segmentation. The following section highlight these phases:

3.1 Data acquisition

The experiment was performed on acrylic photopolymer with 0.8 wt% Boron nitride filler using a clear resin from Formlabs. The sample was printed using Formlabs form2 printer with 25-micron resolution. A NanoTest Platform3 from Micro materials (UK) was used to collect the data using Berkovich Nano indenter. The sample has thickness of 4 mm and diameter of 12 mm. And it was tested for elastic modulus and hardness using load vs. depth hysteresis at 100 indents. 10 x 10 network of test points were set with a 50-micron distance between the points. A maximum load of 20 mN was used with a loading rate of 5 mN/s and an unloading rate of 20 mN/s.

3.2 Denoising

In this research, Jassim (2013) proposed Interquartile Range (IQR) denoised methodology and used it to reduce non-homogeneity values. IQR denoised filter (Figure 1) is implemented by getting a k x k window size, sorting values of that window, calculating the first and third quartiles (Q1 and Q3) and the mean/ average (Q2), finding the thresholds T1 and T2, classifying the values as noise if it is beyond thresholds and averaging the noisy values based on their non-noisy neighbors (Figure 1).

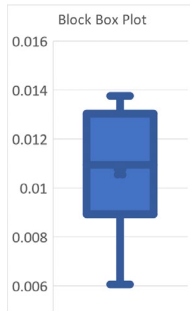


Figure 1. Interquartile Range

In this research two window sizes are tested: 3 x 3 and 5 x 5. Two different approaches are examined to filter the image. First approach is non- overlapping distinct window filtering blocks: once moving over a block of data, it doesn't examine the data partially in next iteration, for example, for a 6 x 6 grid size and 3 x 3 window size, only 4 blocks resulted to apply denoised filter as illustrated in Figure 2.

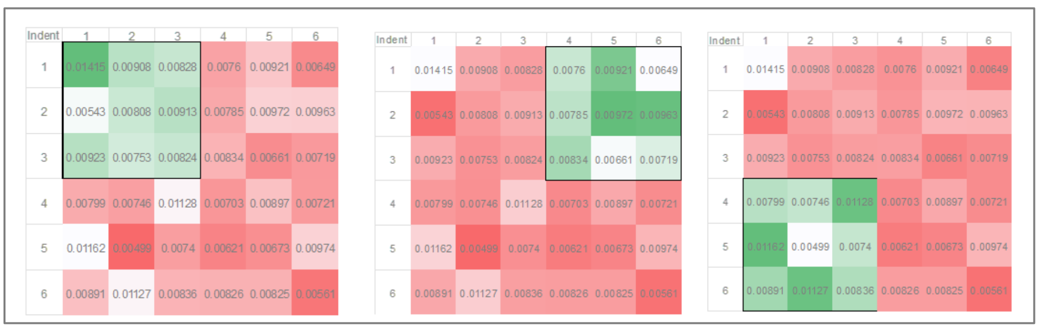


Figure 2. Non- overlapping distinct 3x3 window size, filtering moves steps

The Second approach is Overlapping non-distinct window filtering moves (Figure 3), making intersect blocks to filter, as for 6 x 6 grid size and 3 x 3 window size, 16 blocks resulted to apply filter on.

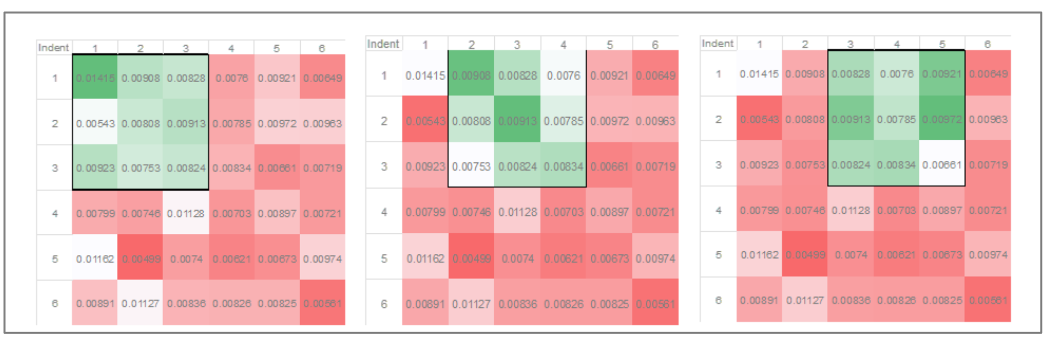


Figure 3. Overlapping non-distinct 3x3 window size, filtering moves steps

Different thresholds ranges calculations were applied: First is wide ranges when lowest threshold T1 was calculated by subtracting Q1 from the difference between Q1 and Q3, and highest threshold T2 was calculated by adding the difference between Q1 and Q3 to Q3 as illustrated in equation 1 and equation 2.

$T1 = Q1 - (Q3-Q1)$ *Equation (1)*
$T2 = Q3 + (Q3-Q1)$ *Equation (2)*

Another calculation approach is proposed in order to narrow the thresholds range and find more noisy data; the narrow ranges where difference was changed to become between Q2 and Q1, Q2 and Q3 as illustrated in equation 3 and equation 4.

$T1 = Q1 - (Q2-Q1)$ *Equation (3)*
$T2 = Q3 + (Q3-Q2)$ *Equation (4)*

3.3 K-means clustering

K-Means Clustering is used after denoising the data. The algorithm was performed using python programming language and Scikit-learn library. Based on Scikit-learn library in python, some of parameters are set by default but the number of clusters has to be set by us. K-Means implementation has few hyper-parameters to consider as shown in Figure 4.

```
class sklearn.cluster.KMeans(n_clusters=8, init='k-means++',
n_init=10, max_iter=300, tol=0.0001, precompute_distances='auto',
verbose=0, random_state=None, copy_x=True, n_jobs=None,
algorithm='auto')
```

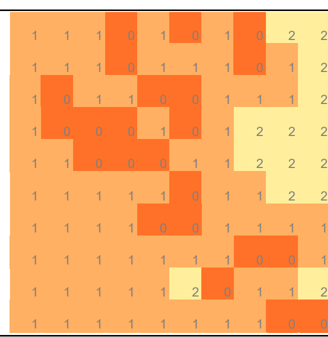
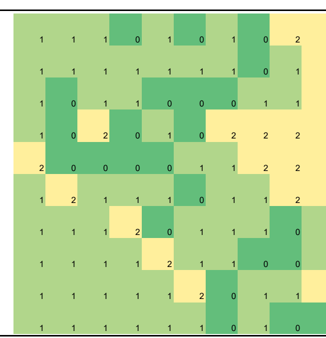
Figure 4. K-Means Hyper-Parameter in Python Syntax

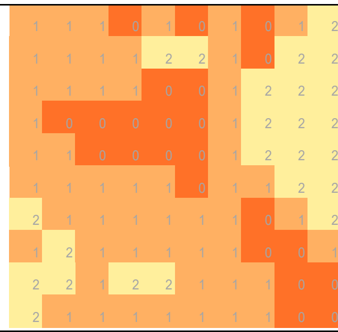
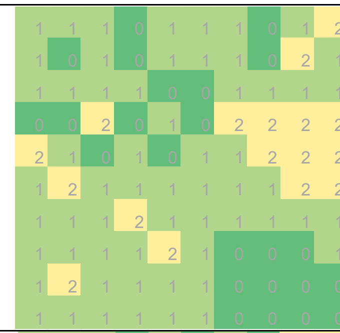
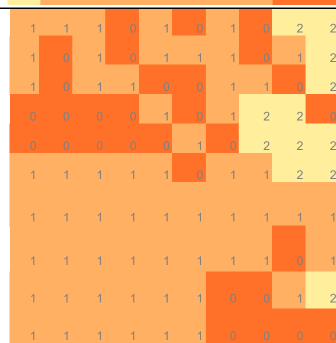
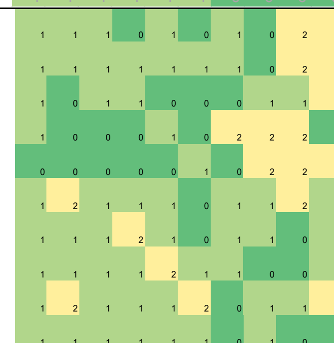
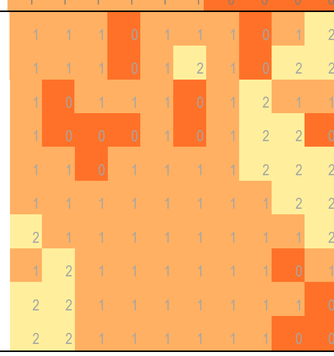
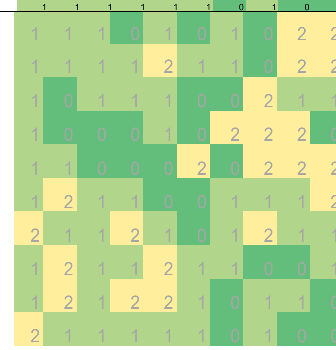
In this implementation, we set number of clusters *n_clusters* to 3, at most as more clusters produce more non-homogeneity areas. Other hyper-parameters kept as the default values. The *n_init* initialize the centroids 10 times and pick the most converging value as the best fit. Increasing this value will scan the entire feature space. The *max_iter* iterates *max_iter* times for each *n_init* runs, i.e., within a run, points will be assigned to different clusters and the loss calculated for *max_iter* times. Keeping *max_iter* at a higher value guarantee that the entire feature space is explored. The *tol* defines relative tolerance of the difference in the cluster centers of two consecutive iterations to declare convergence. After clustering, the colormaps were generated in purpose of visualization.

4. Results and Discussion

Different window sizes were tested with options of overlapping and non-overlapping. With two different threshold ranges were examined; wide range as mentioned in Equation 1 and Equation 2. And narrow range as mentioned earlier in in Equation 3 and Equation 4. The results are illustrated in Table 1.

Table 1. Results of K-Means clusters with different conditions

Size	Threshold Range	Overlapping	Non-Overlapping
Small Window Size (3x3)	Wide (Q3-Q1)		

	Narrow (Q2-Q1) & (Q3-Q2)		
Large Window size (5x5)	Wide (Q3-Q1)		
	Narrow (Q2-Q1) & (Q3-Q2)		

The overlapping de-noise has more values which produced more homogeneous areas. The overlapping, small window size of filtering and narrow threshold range produces more distinct and homogenized areas than wide threshold range. The overlapping and large window size of filtering produces areas that are not distinct regardless the threshold ranges. The segmented areas resulting after non-overlapping produces non homogenous and not distinct areas regardless window size and threshold range. There were values that cannot be denoised automatically, thus further human intervention is required to join points into a cluster. Further investigation is needed to reveal possible factors that could affect the nanoindentation test: the nanoindentation test errors, the grid size, and the mixing methods.

5. Conclusion

A proposed method for segmenting a nano-indentation mechanical property map of a nanocomposite is introduced. Machine learning algorithms were investigated to segment the map into regions with similar mechanical properties. The proposed methodology started with a denoising algorithm to minimize the number of misclassification due to data capturing errors. Different filtering sizes were investigated to achieve the optimal denoising output. K-means clustering algorithm was found to produce distinct clusters with a minimum number of misclassification and human judgment. The proposed methodology can be used in many applications pertaining to material design and characterization.

References

Arshleen, K- Means Clustering Techniques - A Review, *Pramana Research Journal,* vol.8, issue 3, 2018.
Chen, X., Hou, D., Han, Y., Ding, X. and Hua P., Clustering analysis of grid nanoindentation data for cementitious materials*, J Mater Sci*, vol.56, pp. 12238–12255, 2021.

Chmielewski, M., Kaliński , D., Pietrzak ,K. and Włosiński, W., Relationship between mixing conditions and properties of sintered 20AlN/80Cu composite materials. *Archives of Metallurgy and Materials*, vol. 55.2, pp. 579-585, 2010.

Constantinides, G., Ravi Chandran, K.S., Ulm, F.-J. and Van Vliet, K.J., Grid indentation analysis of composite microstructure and mechanics: Principles and validation, *Materials Science and Engineering: A*, vol. 430, Issues 1–2, pp. 189-202, 2006.

Gautham, S. and Saptarshi Sasmal, Recent Advances in Evaluation of intrinsic mechanical properties of cementitious composites using nanoindentation technique, *Construction and Building Materials*, vol. 223, pp. 883-897, 2019.

Goyal, Bhawna, Dogra, Ayush, Agrawal, Sunil, Sohi, B.S. and Sharma, Apoorav, Image denoising review: From classical to state-of-the-art approaches, *Information Fusion*, vol. 55, pp. 220-244, 2020.

Hu, Chuanlin, Nanoindentation as a tool to measure and map mechanical properties of hardened cement pastes. *MRS Communications,* vol. 5, pp. 83–87, 2015.

Hu, Chuanlin and Li, Zongjin, A review on the mechanical properties of cement-based materials measured by nanoindentation, *Construction and Building Materials*, vol. 90, pp. 80-90, 2015.

Jassim, F. A., Image Denoising Using Interquartile Range Filter with Local Averaging, *International Journal of Soft Computing and Engineering (IJSCE),* vol. 2, Issue 6, pp: 424 –428, 2013.

Koumoulos, E., Konstantopoulos, G. and Charitidis, C., Applying Machine Learning to Nanoindentation Data of (Nano-) Enhanced Composites, *Fibers 8*, no. 1: 3, 2020.

Koumoulos, E., Konstantinos P., and Charitidis, C., Constituents Phase Reconstruction through Applied Machine Learning in Nanoindentation Mapping Data of Mortar Surface, *Journal of Composites Science 3*, no. 3: 63, 2019.

Kouqi Liu, Mehdi Ostadhassan, Bailey Bubach, Kegang Ling, Behzad Tokhmechi and Dietrich Robert, Statistical grid nanoindentation analysis to estimate macro-mechanical properties of the Bakken Shale, *Journal of Natural Gas Science and Engineering,* vol. 53, pp. 181-190, 2018.

Menčík, J., Uncertainties and Errors in Nanoindentation, *Nanoindentation in Materials Science*, ch. 3, 2012.

Nagpal, A., Jatain, A. and Gaur, D., Review based on data clustering algorithms, *2013 IEEE Conference on Information & Communication Technologies*, pp. 298-303, 2013

Nohava J, Hausild P, Randall NX. and Favaro G., Grid nanoindentation on multiphase materials for mapping the mechanical properties of complex microstructures*, IMEKO 2010 TC3, TC5 and TC22 Conferences, Metrology in Modern Context*, Pattaya, Chonbu ri, Thailand, November 22 – 25, 2010.

Ronald F. Gibson, A review of recent research on nanoindentation of polymer composites and their constituents, *Composites Science and Technology*, vol. 105, pp. 51-65, 2014.

Syarif I., Prugel-Bennett A. and Wills G., Unsupervised Clustering Approach for Network Anomaly Detection*. In: Benlamri R. (eds) Networked Digital Technologies*. pp. 135 – 145, *Communications in Computer and Information Science*, vol 293, 2012.

Vignesh, B., Oliver, W.C., Kumar, G.S. and Phani, P.S., Critical assessment of high speed nanoindentation mapping technique and data deconvolution on thermal barrier coatings, *Materials & Design,* vol. 181, 2019.

Zhiyu Luo, Wengui Li, Yixiang Gan, Kavya Mendu and Surendra P. Shah, Applying grid nanoindentation and maximum likelihood estimation for N-A-S-H gel in geopolymer paste: Investigation and discussion, *Cement and Concrete Research*, vol. 135, 2020.

Biographies

Sarah Ghaleb Alhamdani is a student in Master of Engineering and System Management at Alfaisal University. She received her BSc in Computer Science from Umm-AlQura University, Makkah in 2015.

Abdallah Alrshdan is an assistant professor in Industrial Engineering at Alfaisal University, Saudi Arabia. He teaches courses in Ergonomics, Work Design, Data analytics, and Quality Engineering. His PhD in Industrial Engineering is from Wichita State University. His current research focuses on ergonomics product design, applications of AI in the design and manufacturing process, and lean production. He worked as a production manager at ALL Cell technologies in USA building Li-ion batteries used for electrical cars and continues as a consultant in the research and development department.

Dr. Alrshdan is currently the chair of the Industrial Engineering department and the head of quality assurance. He serves as a reviewer for different international journals and session chair in international Industrial Engineering conferences.

Khalid Ahmed is a lecturer in College of Engineering at Alfaisal University. He revieved his Master Degree in Polymer Science and Technology. Experience in process industries, teaching engineering courses and labs**.** Background in polymeric materials and research interest in SLA 3D printing and phase change materials.