ARL-TR-9532 ● SEP 2022

DEVCOM ARMY RESEARCH LABORATORY

Adaptive Data Analysis for Fatigue Studies of Additively Manufactured 17-4 PH Stainless Steel

by Natasha C Bradley, Michael D Coatney, Todd C Henry, and Mulugeta A Haile

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TR-9532 ● SEP 2022

Adaptive Data Analysis for Fatigue Studies of Additively Manufactured 17-4 PH Stainless Steel

Natasha C Bradley, Todd C Henry, and Mulugeta A Haile
DEVCOM Army Research Laboratory

Michael D Coatney
Army Test and Evaluation Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
September 2022	Technical Report	October 2021–May 2022

4. TITLE AND SUBTITLE
Adaptive Data Analysis for Fatigue Studies of Additively Manufactured 17-4 PH Stainless Steel

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Natasha C Bradley, Michael D Coatney, Todd C Henry, and Mulugeta A Haile

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DEVCOM Army Research Laboratory
ATTN: FCDD-RLW-VB
Aberdeen Proving Ground, MD 21005

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-9532

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES
ORCID IDs: Mulugeta Haile, 0000-0002-3038-9155; Todd Henry, 0000-0002-3378-3780

14. ABSTRACT
In this report, principal component analysis (PCA) is used to investigate quasi-static and tension–tension fatigue stress and strain data sets of modified ASTM E8 subsized dog-bone-shaped samples manufactured by a Markforged Metal X additive manufacturing printer. These large data sets were made using the strain field data created using digital image correlation (DIC) imaging technology. Using PCA unsupervised learning algorithms, we investigated if these machine learning techniques can provide various trends and patterns of these materials that DIC and other traditional material characterization techniques cannot show us. Although a simple PCA of images did not show a trend or unique pattern within the data, it did, however, show that reducing the data size and processing via DIC does not cause a large reduction in the DIC strain output.

15. SUBJECT TERMS
Mechanical Sciences, digital image correlation, principal component analysis, machine learning, additive manufacturing, 3-D printed metal

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Natasha Bradley
Unclassified	Unclassified	Unclassified	UU	19	19b. TELEPHONE NUMBER (Include area code) (410) 278-7756

Contents

List of Figures

1. Introduction

Machine learning (ML) and artificial intelligence (AI) has changed the way we use data and found ways for us to understand trends and patterns that would not be possible with traditional analytical methods.[1] ML can effectively and efficiently identify patterns in large high-dimensional data sets, extract desired information, and discover new laws. It is widely used in analyzing material property (degradation detection, nanomaterials analysis, and molecular property prediction)[2,3] and discovering new materials (structure-oriented design, element-oriented design, inverse design, drug design, and quantum chemistry).[4] ML's principal component analysis, or PCA, is one of the more popular techniques used today. PCA is a multivariate technique that analyzes a data table in which observations are described by several intercorrelated quantitative dependent variables.[5,6] Also, PCA is a technique for reducing the dimensionality of such data sets, increasing the interpretability but at the same time minimizing information loss.[7] After a high-dimensional graph is created, a low-dimensional graph is constructed using fewer dimensions than that provided by the complex data set data and it tries to replicate the high-dimensional graph.

Additive manufacturing (AM) of metals is one of the most recent manufacturing techniques of the past few decades. AM is associated with the layer-by-layer buildup of structures with materials such as thermoplastic polymers, metals, ceramics, and more recently, fiber-reinforced composites.[8] As a result, this reduces waste and allows businesses to produce parts faster and at a lower cost than traditional subtractive techniques. AM methods can be used to print intricate structures where growing complexity increases the AM's attractiveness over conventional manufacturing for design processes like topology optimization.[9,10] Conventional manufacturing methods will have much greater strain to failure and typically are also stronger and stiffer depending on the way the part is printed.[11] Metal AM parts are sensitive to microstructure where parameterization of things like laser power and speed can change performance.[12–16] Non-laser-based methods extrude a highly filled wax-polymer binder for which the wax is later removed in a washing step and the part consolidated via sintering. This process is called atomic diffusion AM (ADAM). ADAM parts exhibit some of the same defect structures as similar polymer fabrication methods for which mechanical performance is typically worse than laser/powder methods but arguably safer.[17–20]

The work presented here discusses the PCA ML algorithms as applied to experimental data collected for ADAM processed test samples. Test data includes digital image correlation (DIC) images and complex strain data sets of several quasi-static and tension–tension fatigue specimens. The digital image PCA results

will focus on the replication of images that give DIC results as similar to that of the original DIC results. The PCA on the observed strain data will look for any trends or patterns present, seen by ML algorithms that cannot be seen using just stress–strain relations as a function of loads and displacements.

2. Experimental Data for ML

A MarkForged Metal X printer was used in this research to create subsized ASTM E8[21] samples using the ADAM process. Sample geometry had an approximate length, width, and thickness of 70, 6.2, and 1.6 mm, respectively, with a 1.7-mm-circular hole centered in the width. The stress concentration factor at the hole was therefore around 2.37. Surface strains were measured using DIC, which calculates full-field strains noncontact via stereo cameras.[22,23] Two 5-megapixel cameras were used to capture and process DIC data. Mechanical testing was conducted using an MTS hydraulic test frame applying either 2 mm/min quasi-static loading or force-controlled fatigue loading at 5 Hz and an R ratio of 0.1. Force, displacement, and time information was recorded by the test frame (Fig. 1).

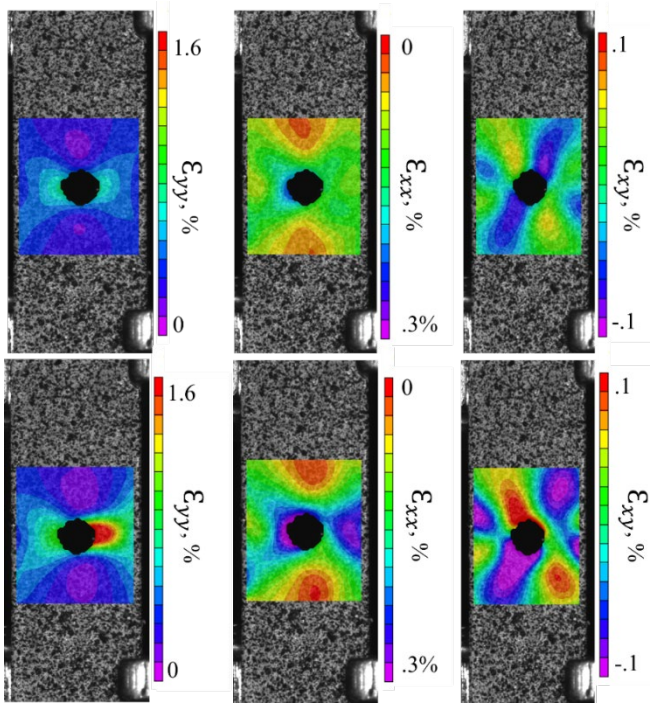

Fig. 1 Example axial (ε_{yy}), transverse (ε_{xx}), and shear (ε_{xy}) strain as measured by DIC for a sample at the start of loading (top) and near failure (bottom)

2.1 Unsupervised ML

The goal of unsupervised ML is to identify or classify the data collected into groups that are similar. This type of learning has some important characteristics that could be extremely useful in data analysis, especially using large sets of data. One characteristic is that unsupervised learning can be performed on data with or without labels, which is critical because data sets can have incomplete or missing information. Another characteristic is that unsupervised learning can be done real time, meaning it can classify on the spot without waiting for the machine to make a complete inference from the training data. Unsupervised learning can also make connections within data that has been previously unknown because all it is looking for is a pattern or similarity. This can be particularly useful when exploring new materials or a combination of materials, where the similarities and differences to parent materials may be hard to quantify.

2.2 Preprocessing

To investigate the use of an unsupervised model, first the data needed to be preprocessed to prepare the data for the model. There were four steps taken to preprocess the data, as seen in the Fig 2. Each of these steps helps clean the data and make it more suitable for ML analysis. These steps also improve accuracy and efficiency of the analysis.

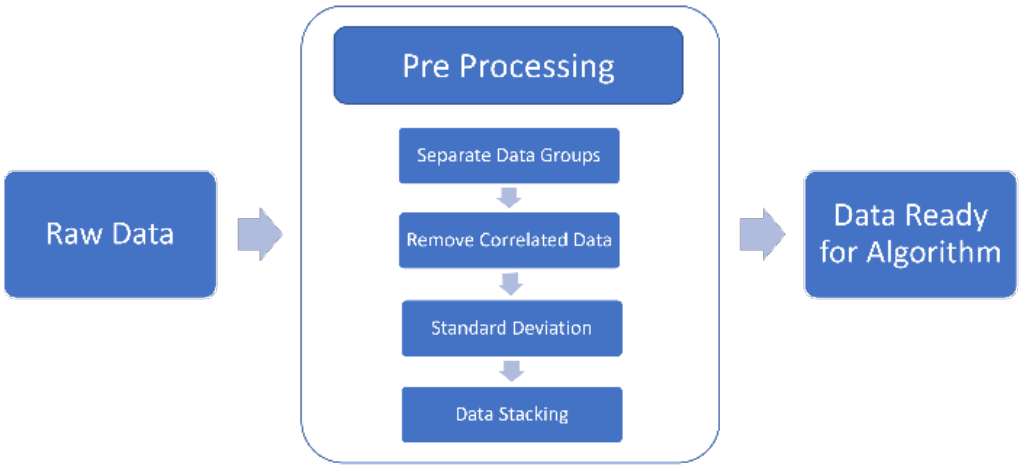

Fig. 2 Preprocessing method before data is submitted to the ML algorithm

Step 1 – Separate Data Groups: The data collected was DIC data for three different points on a given sample: 1) near the hole right-side, 2) near the hole top, and 3) away from the hole. Although it can be analyzed as a group across a sample, it can also be analyzed as a specific point that is in the same place across multiple samples. Another separation that had to occur is when performing the PCA on

images, the left camera image is different from the right camera image because the perspective from each camera is different.

Step 2 – Remove Correlated Data: Much of the data collected from the DIC is length, position, and displacement data. These data are often similar to one another and sometimes can be exactly the same because the sample is not very large before failure. The other issue is that because quasi-static images are acquired at 1 Hz (fatigue data is collected for every peak and valley) and sample failure is much faster, evolution of the moments immediately before failure are lost. This means that position and displacement data may only change in one or two data lines. Therefore, this data is removed before we begin our analysis.

Step 3 – Standard Deviation: Specifically, in the PCA the standard deviation is taken to standardize the data in reference to how it compares to itself. Taking the standard deviation tells how each point is compared to the center of the data.

Step 4 – Data Stacking: The last step of the preprocessing was stacking the data into large groups based on the point the strain was taken. So, when analysis is performed it can be performed on multiple tests that are similar. This helps with labeling and finding more similarities by comparison, such as: is each type of test showing the same patterns? Or is there a pattern between multiple test specimens under similar conditions?

2.3 PCA

PCA is a technique for reducing the dimensionality of data sets—exploiting the fact that the images have something in common. Images are large files with a grid of pixels. For each picture you can have 5,013,504 dimensions before any cropping or PCA is performed. Large numbers of dimensions make it difficult to visualize similarities because the groups would be in dimensional spaces that are impossible to visualize; for example, a 5-D plot. PCA uses covariance and eigenvectors to effectively reduce the dimensions with little losses. The PCA can also show the number of vital dimensions for image reconstruction. The PCA uses eigen values and vectors to determine principal components (PCs) from the covariance matrix. These PCs are combinations of the features (i.e., for images each pixel is a feature) that are uncorrelated from one another but carry the most information possible from the remaining information minus the previous PC.

PCA can also be used for the strain data collected from the DIC image. This data contains six dimensions—one dimension for each line of data collected or calculated by the DIC process. Although this is significantly fewer dimensions than the image data, reduction is still valuable because you cannot visualize a 6-D plot.

Both of these different PCAs were conducted to determine if the data reduction would produce a trend or some visually noticeable change in the data.

3. Results

3.1 PCA on Images

The process used to perform the PCA on the image data was to first compute its intrinsic dimensionality, then plot the relative value of the information content of each of the PCs and compare them (Fig. 3). Then a scree plot was created to determine the number of factors that dominate the space. This plot, shown in Fig. 3, shows the scree plot. When using images, there are only a few PCs needed to reduce the information collected from the eigenvalues. The scree plot and PCA plot shown in Fig. 3 indicate that a dimension reduction from 140 components to 3 components will still result in more than 97% of the information captured within each image.

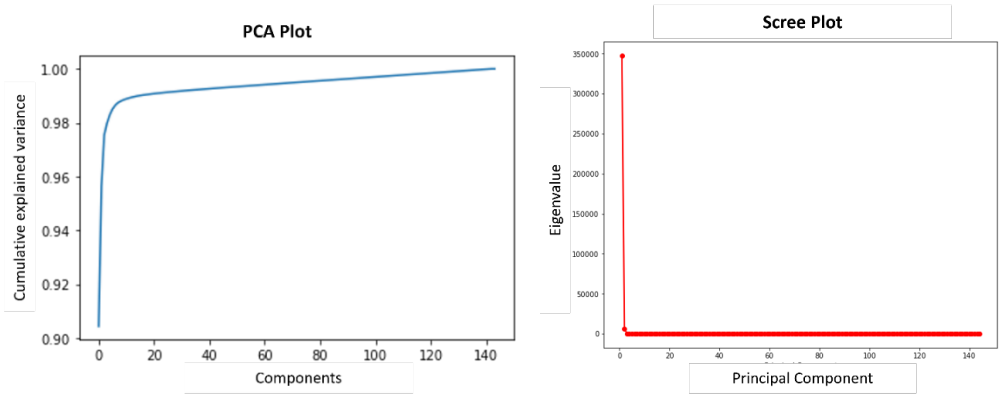

Fig. 3 PCA plot and scree plot of a single set of DIC images for a quasi-static tension test

This dimension reduction can be calculated by looking at the proportion of variance found in each PC and adding it to give a cumulative explained variance. The proportion of variance is simply a function of ratio of related eigenvalue (λ_{PCi}) to that PC and sum of eigenvalues of all eigenvectors ($\lambda_{PC1}+\lambda_{PC2}+\ldots+\lambda_{PCn}$), given i is the related eigenvalue in n total PCs. Eq. 1 gives the calculation for the proportion of variance using the PCA method.

$$Proportion\ of\ Variance = \frac{\lambda_{PCi}}{\lambda_{PCn}+\lambda_{PC2}+\cdots+\lambda_{PCn}} \quad (1)$$

Once the number of PCs needed was determined, the data was then reconstructed into its original form to feed back into the DIC system for analysis. This would normally cause a reduction in file size; however, because the DIC system uses a

layered Tag Image File Format (TIFF) image, the size of all TIFF images are the same. This new TIFF contained less data with only 59,675 dimensions, more than 99% of the dimensions reduced. As seen in Fig. 4, when the data is reconstructed back into an image, the quality reduction is slight. Because the PCA uses just information from the image only and no actual data, the concern was if there would be any shift in the information obtained by inputting this information back into the DIC and running the analysis again to obtain the strain fields as mentioned earlier in this report. Figure 5 shows some results from this analysis. The results were very similar, specifically for the first 10 images; after that, the in-plane strain recorded seems to be slightly smaller than the actual strain of the unaltered images.

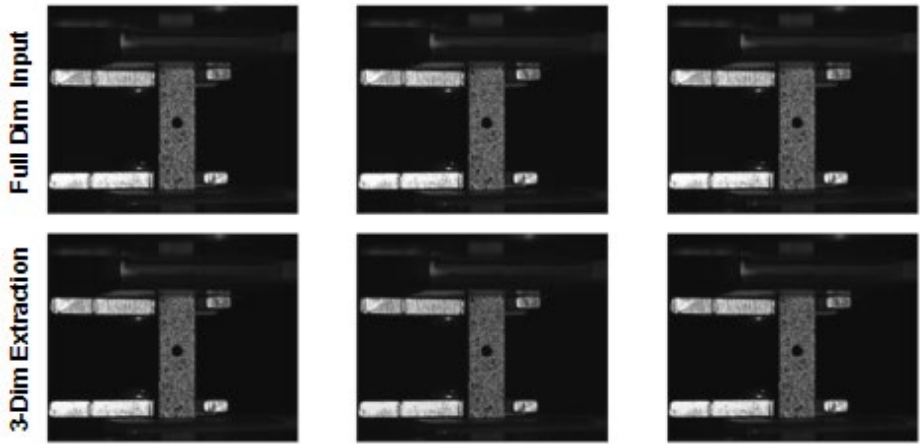

Fig. 4 Comparison of PCA original images vs. 3-D PCA reconstructed images

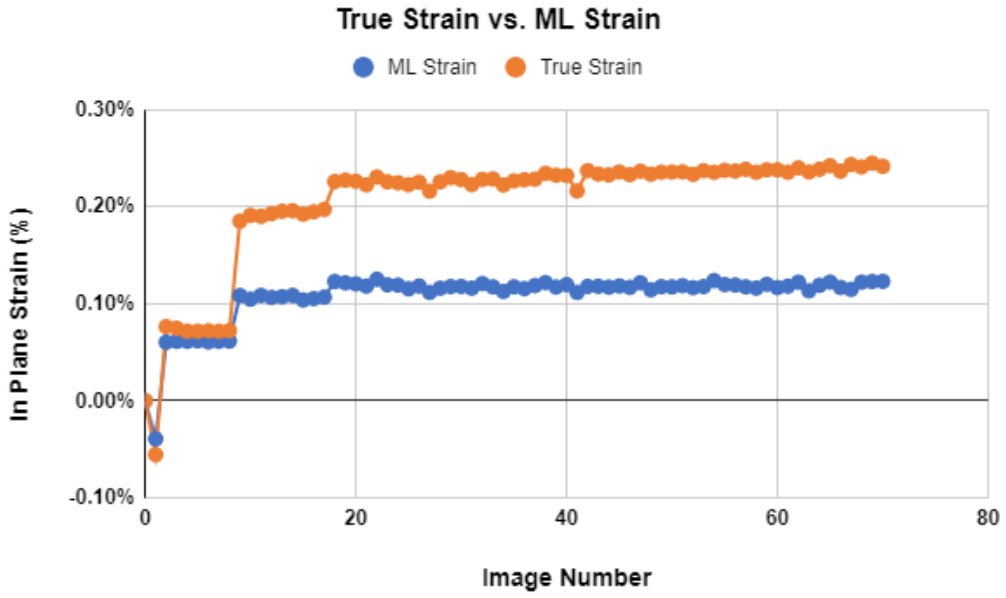

Fig. 5 Strain data comparison of PCA original images vs. 3-D reconstructed PCA image results

3.2 PCA on Strain Data

Instead of a PCA on the image as seen in Fig. 5, the following are the results of a PCA done with DIC output data only on some fatigue tests and quasi-static tests done with the same parameters. This data used the strain columns including the strain in the x direction, the y direction, the shear strain, the major principal strain, the minor principal strain, and the gamma, which is the principal strain angle. All displacement data collected was discarded for this test. Negative strain and positive strain is only relevant based on the axis direction established when testing, so none of the strain values were changed before this analysis was performed. The PCA on the data was performed to specifically reduce the number of components needed to understand the data from six to two. This is because it is significantly easier to understand a 2-D plot than a 6-D plot. When the cumulative variance was calculated, 90% of the information was collected in the first two PCs. The loadings calculated for each test type are similar, meaning the first PC was affected most by the strain data in the x and y, as well as principal strains, and the second PC is mostly affected by the strain angle and the shear strain. The loadings are a linear combination of the original variables from which the PCs were constructed.

The first analysis was done on six quasi-static samples. The results were separated and then labels were put on the first and last data points in the cycle. In the results shown in Fig. 6, you can see that the data does not perform clusters at all but seems to increase linearly until the first and last cycles are separated out. It is noticeable that the first and the last cycles are more affected by PC 1 than PC 2. The results for the quasi-static test were similar in that they yielded similar ultimate strength (UTS) numbers. However, when looking at how the strain seen in each sample is affected throughout the test, we see that the samples vary greatly—specifically Samples 2 and 4 have a higher dependency on Component 2 than the other samples. The slope of the scatter of each sample's data is close to the same—this could be related to the UTS being closer than the relationships of PCs 1 and 2 show in the analysis.

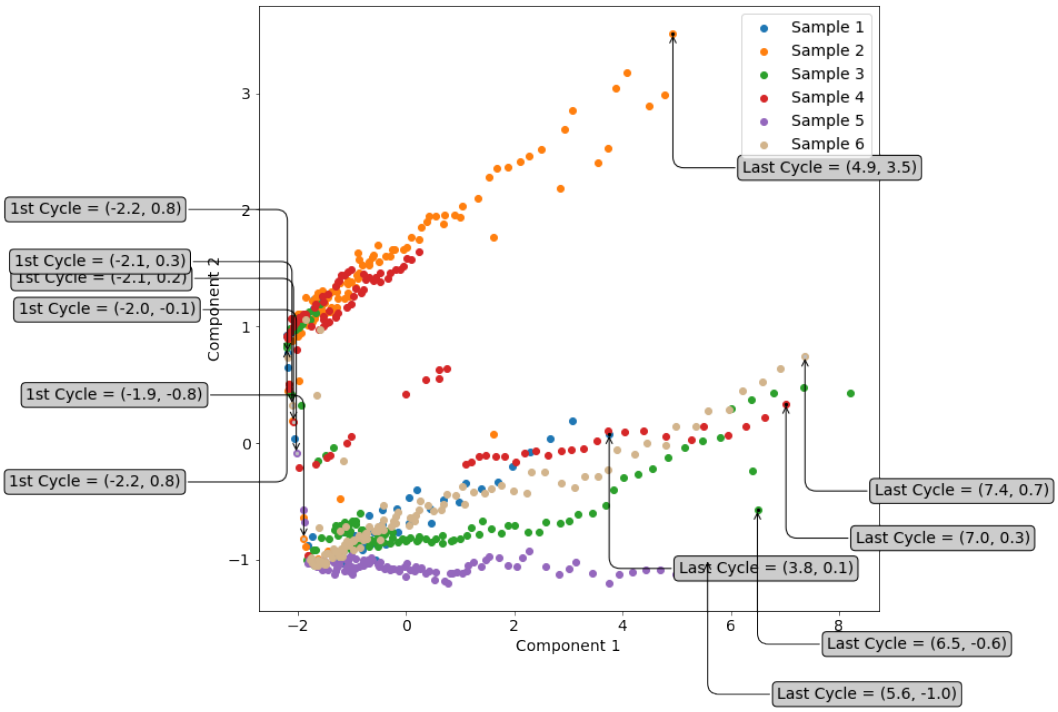

Fig. 6 PCA results on strain data for quasi-static tension test; first and last cycle of each sample is labeled

The second analysis was done on four different fatigue samples, and then the results were separated and labels were put on the first and last data points in the cycle. In the results shown in Fig. 7, you can see that the data for Samples 1, 3, and 4 is separated into two different clusters. It is also easy to notice that the first and the last cycles are never in the same cluster of data points. However, if you look at the colored arrows, the arrow for Sample 4 is decreasing as the relationship between PC 1 and PC 2 decreases, whereas the other samples show an increasing relationship. It would be interesting to investigate why this occurs as the data does not show anything abnormal during that test sample.

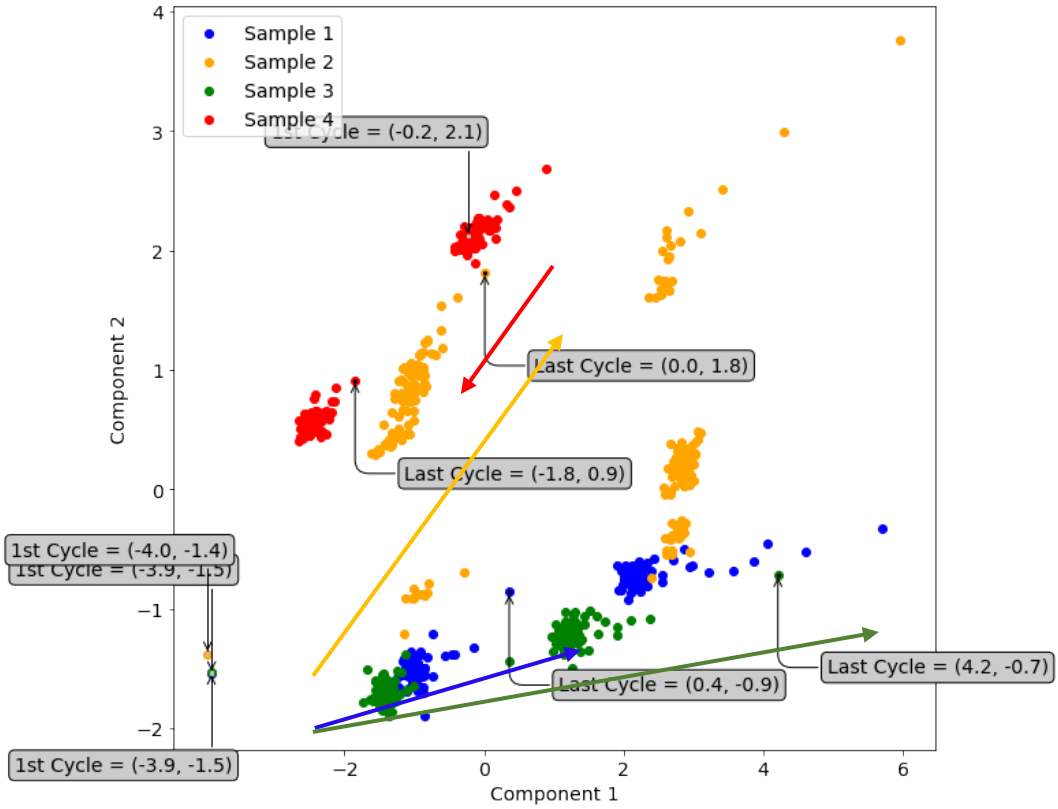

Fig. 7 PCA results on strain data for tension–tension fatigue test; first and last cycle of each sample is labeled

4. Conclusion

As the testing results show, the removal of information by the PCA and reduction of image quality does cause a slight change in the DIC's results. The fact that these reconstructed images can be used tells us that if there is corrupted data, or data that is not in the correct format for the DIC system, it can be easily converted to the correct format and then submitted with very little deviation from the desired results. It would be worth further investigating if any image can be converted into a DIC-ready format as well as how much the image can be compressed to produce a similar result to the original.

5. References

1. Alloghani M, Al-Jumeily D, Mustafina J, Hussain A, Aljaaf AJ. A systematic review on supervised and unsupervised machine learning algorithms for data science. In: Berry M, Mohamed A, Yap BW, editors. Unsupervised and semi-supervised learning. Springer International Publishing; c2019. p. 3–21.

2. Bock FE, Aydin RC, Cyron CJ, Huber N, Kalidindi SR, Klusemann B. A review of the application of machine learning and data mining approaches in continuum materials mechanics. Front Mater. 2019;6:110.

3. Cecen A, Dai H, Yabansu YC, Kalidindi SR, Le S. Material structure-property linkages using three-dimensional convolutional neural networks. Acta Mater. 2018;146:76–84.

4. Abdi H. Principal component analysis. WIREs Comp Stats. 2010;2(4):433–459.

5. Wei J, Chu X, Sun X, Xu K, Deng H, Chen J, Wei Z, Lei M. Machine learning in materials science. InfoMat. 2019;1:338–358.

6. Abdi H, Williams LJ. Principal component analysis. In: Li SZ, Jain A, editors. Encyclopedia of Biometrics. Springer; 2009.

7. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. Philos Trans A Math Phys Eng Sci. 2016;374(2065):20150202.

8. Razavykia A, Brusa E, Delprete C, Yavari R. An overview of additive manufacturing technologies—a review to technical synthesis in numerical study of selective laser melting. Materials. 2020;13(17):3895.

9. Vanderesse N, Richter A, Nuno N, Bocher P. Measurement of deformation heterogeneities in additively manufactured lattice materials by digital image correlation: strain maps analysis and reliability assessment. J Mech Behav Biomed Mater. 2018;86:397–408.

10. Henry TC, Johnson TE, Haynes RA, Tran A. Fatigue performance of polyamide 12 additively manufactured structures designed with topology optimization. J Test Eval. 2021;49(3):1–19.

11. Cole DP, Gardea F, Henry TC, Seppala JE, Garboczi EJ, Migler KD, Shumeyko CM, Westrich JR, Orski SV, Gair JL. AMB2018-03: Benchmark physical property measurements for material extrusion additive manufacturing of polycarbonate. Integr Mater Manuf. 2020;9:358–375

12. Godec D, Cano S, Holzer C, and Gonzalez-Guiterrez J. Optimization of the 3D printing parameters for tensile properties of specimens produced by fused filament fabrication of 17-4 PH stainless steel. Materials. 2020;13(3):1–23.

13. Agrawal P, Haridas RS, Thapliyal S, Yadav S, Mishra RS, McWilliams VA, Cho KC. Metastable high entropy alloys: an excellent defect tolerant material for additive manufacturing. Mater Sci Eng A. 2021;826:142005.

14. Shu Y, Galles D, Tertuiliano OA, McWilliams BA, Yang N, Cai W, Lew AJ. A critical look at the prediction of the temperature field around a laser-induced melt pool on metallic substrates. Sci Rep. 2021;11(1):12224.

15. Beretta S, Romano S. A comparison of fatigue strength sensitivity to defects for materials manufactured by AM or traditional processes. Int J Fatigue. 2017;94:178–191.

16. Flodberg G, Pettersson H, Yang L. Pore analysis and mechanical performance of selective laser sintered objects. Addit Manuf. 2018;24:307–315.

17. Gonzalez-Gutierrez J, Cano S, Schuschnigg S, Kukla C, Sapkota J, Holzer C. Additive manufacturing of metallic and ceramic components by the material extrusion of highly-filled polymers: a review and future perspectives. Materials. 2018;11(5):840.

18. Wu G, Langrana NA, Rangarajan S, McCuiston R, Sadanji R, Danforth S. Solid freeform fabrication of metal components using fused deposition of metals. Mater Des. 2002;23(1):775–782.

19. Galati M, Minetola P. Analysis of density, roughness, and accuracy of the atomic diffusion additive manufacturing (ADAM) process for metal parts. Materials. 2019;12(24):1–15.

20. Suwanpreecha C, Manonukul A. On the orientation effect in as-printed and as-sintered bending properties of 17-4 PH alloy fabricated by metal fused filament fabrication. Rapid Prototyp J. 2022;28(6):1076–1085.

21. ASTM E8. Standard test methods for tension testing of metallic materials. ASTM International; 2022.

22. Grédiac M, Hild F, editors. Full-field measurements and identification in solid mechanics. John Wiley & Sons; 2013.

23. Sutton MA. Digital image correlation for shape and deformation measurements. In: Springer handbook of experimental solid mechanics. Springer; c2008. p. 565–600.

List of Symbols, Abbreviations, and Acronyms

2-D	two-dimensional
3-D	three-dimensional
5-D	five-dimensional
6-D	six-dimensional
ADAM	atomic diffusion AM
AI	artificial intelligence
AM	additive manufacturing
DIC	digital image correlation
ML	machine learning
PC	principal component
PCA	principal component analysis
TIFF	Tag Image File Format
UTS	ultimate strength

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	DEVCOM ARL FCDD RLD DCI TECH LIB
5 (PDF)	DEVCOM ARL FCDD RLW VB N BRADLEY M HAILE A HALL T HENRY FCDD RLW V S SILTON
1 (PDF)	ATEC M COATNEY