This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

nature portfolio

Peer Review File

Predicting Compressive Stress-Strain Behavior of Elasto-Plastic Porous Media via Morphology-Informed Neural Networks

Corresponding Author: Dr Winston lindqwister

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The proposed manuscript proposes a procedure to calibrate a neural network to deduce, from microstructural descriptors, the nominal stress-strain curve of a compressive test on a cubic sample. The topic is interesting and a lot of work has been done, however it is the Reviewer's opinion that some issues still exist as listed below:
- please double check the manuscript, some typos still remain.
- the selection of the REV volume seems a bit arbitrary (pg. 7). For a ratio of 0.1 significant scattering can be seen in fig. 2. Please better justify the adopted choice.
- it is not clear how the mesh is created. The creation of a mesh of a porous material from voxel data is not trivial and might lead to error if the features are very close. How do the Authors controlled the mesh generation? Is it a voxel-based mesh? Which finite elements have been adopted? Do the Authors noticed any locking effects due to plasticity?
- many doubts remain also for the network training. Please explain better the procedure that has been followed, especially the stopping criterion used. Also, this type of problems are often non-convex and many sets of parameters might lead to low cost functions. Has this been taken into account? How do the Authors address this problem?
- in general, it is not completely clear what is the scope and the main goal of this work. As proposed, it allows to surrogate the determination of stress-strain pairs for a compressive test, however, for a comprehensive characterization of a material much more information are needed, such as tensile behavior, permeability, flexural behavior to name a few. Also, the choice of doing compressive tests on cubic samples is not optimal for, e.g., the Young's modulus due to the edge effects. Although the proposed work presents a methodology rather than an approach for a specific material, it is not clear what the breath of the work is. Do the Authors envision the adoption of this type of approach for a specific material? If not, what is the main aim of this method?
- the Authors also limit the study to elasto-plastic behavior, this can be a limiting assumption since damage might take place as well. Also, contact between different parts of the microstructure can take place.
- can the Authors comment on the expected precision of the CT scan images? This can limit the generality of the descriptors if some specific features are not captured.
- did the Authors try an inverse problem where given a stress-strain curve a set of descriptors is given? This might be an interesting extension to a-posteriori estimate the descriptors from experimental tests.

Reviewer #2

(Remarks to the Author)
This paper presents a machine learning (ML) approach to predict the effective stress-strain response of porous materials. Experimental tests and FE simulations are conducted to extract the target responses. The proposed ML methodology then recalls Hadwiger's theorem to select four morphological parameters as input features. It is shown numerically that this reduced parameter space yields accurate results, almost as accurate as an alternative model that involves 31 additional parameters.

The paper is well-written and easy to follow. Moreover, the results highlight the ability of the ML approach to predict the effective stress-strain response of porous materials with a reduced set of parameters. However, while the ideas can potentially represent a worthy contribution to the journal, the paper presents some shortcomings that should be addressed to be considered suitable for publication. I list my main concerns below.

1) The novelty is not sufficiently clear. What is the benefit of this approach vs. other (rather abundant) works in the literature

that also rely on ML to predict global stress-strain responses?

2) I struggle to see the added value of an ML approach for reproducing the global stress-strain behavior. In particular, if the only goal is to capture the effective stress-strain behavior, one may use simple empirical models or constitutive plasticity theory at the material point level to yield the target response without resorting to FEA. Of course, for more ambitious objectives, e.g., predicting full-field quantities or averaged properties that require full-field microstructural solutions, one could argue that FEA simulations are expensive and that ML approaches may come in handy, but this does not seem to be the case here. What is the added value of the present ML approach in this context?

Minor comments

3) The description of the geometric parameters is a little vague. I suggest adding a few details. For instance, how is the mean curvature captured by the grain size? How is the mean breadth computed if not all solid domains are convex?

4) Why is the J2 plasticity model chosen? Again, this point may reflect the fact that only the global response is studied, which can be captured in many ways.

5) Simulations with finer meshes could be conducted to confirm that the results of poorly predicted microstructures are due to over-coarsening.

6) Can the authors comment on the dependence on curing time and temperature, particularly considering that only morphological parameters are inputs?

7) At the beginning of page 8: did the authors mean "exponentially [decrease] computational time"?

Reviewer #3

(Remarks to the Author)
Summary
The submitted manuscript proposes a machine learning approach predicting stress-strain relationships for microstructural geometries. To this end, extracted geometric features of the underlying microstructure are mapped with a deep Neural Network (NN) to stress-strain curve corresponding to a uniaxial test. Of specific interest is a sub-class of descriptors consisting in Minkowski functionals, representing volume, surface area, mean curvature, and total Gaussian curvature of the microstructural grains. It is shown that the set of four Minkowski functionals constitutes an appropriate/effective set of descriptors, as on their basis high accuracy in predictions can be achieved (R-square value of 0.939) as compared to a full set of used geometrical features (R-square value of 0.943) for the microstructures considered. For data generation, numerical simulations were carried out in a FEM package MOOSE, which were used to identify material properties by comparison against experimental observations of manufactured specimens.

Merits
This is an interesting contribution. Its main originality lies, in my opinion, in utilizing Minkowski functionals for predicting stress-strain curves of complex 3D microstructures in combination with deep NNs.

Critique
I have, nevertheless, the following three main questions/concerns, which I would invite the authors to comment on and clarify:
1. It is not clear from the manuscript whether performance of multiple feature combinations have been tested in place of microstructural descriptors. The authors mention on page 10 that “multi-dimensional search grid of 200 parameter combinations was developed”, but later comment that “the model was retrained with only the four Minkowski functionals”, which makes it somewhat inconclusive.
If other feature combinations have not been tested, then the main statement of the manuscript is in my opinion not substantiated, since the authors do not convincingly show/demonstrate performance of other feature combinations (e.g., three instead of four Minkowski functionals, or other combinations of geometrical descriptors or completely different descriptors, cf. comment 20 below). This would then weaken the main hypothesis of the manuscript and would make the entire study fairly inconclusive.
2. The role of experiments is not clear. The authors seem to employ experiments to verify numerical predictions against experimental measurements (or, vice versa), and to identify material parameters of the bulk material model. This seems to be unnecessary since correlation between geometrical features and stress-strain relationships through NN is being tested and any ad-hoc material (with plasticity) would serve the purpose equally well. The study and the main results do not seem to benefit from the conducted experiments, and even seem to cloud the main message of the submitted manuscript.
3. The structure of the paper is not very clear and resembles a report rather than a manuscript with a carefully build argumentation; see also comment 15 below.

Technical Comments
Furthermore, I have also a few technical comments:
4. Page 4: “Minkovski functionals can fully describe a microstructure”. Certainly, a too strong statement, in addition contradictory with latter statement on page 12 “…it was hypothesized that these four functionals…”. Please clarify and correct.

5. Introduction: it would be good to clearly specify the considered setup early on in the manuscript and mention that only uniaxial tests are considered, leading to only one component of stress-strain relationships.
6. Section B: representative elementary volume. How does this relate to typically used representative volume element (RVE), known from computational homogenization?
7. Section C: "…the image was binarized to separate the solid…"; how exactly has this been done, what was the threshold value, etc.? Similar details may significantly influence the results. Additional question: could the authors elaborate on how the proposed methodology would work for multi-material systems? Binary composites considered, i.e., bulk/void, seem to be fairly limited for practical applications.
8. Page 7: "motion-invariant—or isotropic—microstructure". Could the authors specify more what they mean with isotropic? There seems to be a collision in terminology (the text seems to discuss group of transformations, while isotropic microstructures can be interpreted also as those that result in isotropic effective material properties/behaviour).
9. Page 7/8: "To speed up the collection … to potentially reduce the order of nodes on the mesh and to exponentially increase computational time …" decrease computational time?
10. Page 8: "finite element results were compared to physical experiments": the role of experiments is not clear in the manuscript, cf. point 1. Please clarify.
11. Table I: entries in the last two rows and first three columns seem to be missing.
12. Table II: again, entries missing and not clear how to interpret the blank space in the table.
13. Section B: "experimental validation of FEA" not clear why this section is present and why FEA needs to be validated.
14. Page 12: "scaling the size of the printed cube had a minimal effect". Is this assuming fixed geometry? If yes, then this exercise confirms scale-independence of the employed constitutive law, which is clearly to be expected a priori. The follow-up statement "This supports the assumption … is representative" seems to be then wrong. Please clarify and elaborate.
15. Page 13: "the overall goal for the project", "The results of this project" seem to suggest a report-like approach towards the submitted manuscript; please refer to comment 3 above and correct/clarify.
16. Page 13: "demonstrate not only the capabilities of Minkowski functionals…". This statement does not seem to be substantiated properly, since performance of Minkowski functionals has not been compared against other descriptors. It can be hypothesized that deep NN could compensate for the effect of using other descriptors. See also comment 1 above.
17. Page 14: "The minimal reduction in performance of the model when using only the four Minkowski functionals demonstrates…"; cf. comment 16. The claim is not well substantiated.
18. Page 15: "the focus of this work was to link theory, simulation, experiment"; the role of either simulation or experiment is not clear. Why both and not, for instance, numerical simulations only?
19. Conclusions: could the authors comment on the accuracy of predictions in terms of different loading conditions (assume, for instance, a typical RVE element in computational homogenization) and not a single tensile test, employing all deformation modes. To which extend are achieved conclusions transferrable to such situations?
20. Appendix B: It is not clear to the reviewer what was the motivation for including the listed descriptors. Why not including standard descriptors corresponding to random microstructures (see, e.g., Torquato, Random heterogeneous materials, for more details).

Details
Finally, a few minor textual details:
15. p. 12: "the reactive force" → "the reaction force"; please correct elsewhere.
16. …

Discussion
Although the topic covered in the submitted manuscript is interesting and fits thematically well within the scope of the journal, based on the arguments outlined above (mainly points 1-3) I unfortunately cannot recommend it for publication in its current form and invite the authors to address the above issues in a major revision.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Most of the remarks raised have been addressed, however it is still not clear how this approach should be able to improve discovery of new materials (reply to point 5 raised) only considering the compressive behavior. There is no evidence that this is enough, considering that the method does not take into account any other mechanical behavior or characteristic than the compressive stress-strain curve. The Reviewer suggests to reduce the breath of the method to those application where only the compressive behavior characterized by J2 plasticity is needed/relevant.

Reviewer #2

(Remarks to the Author)
The authors provide, for the most part, a satisfactory response to the remarks raised in the reviews. However, some crucial aspects are still not sufficiently convincing. I list my main concerns below.

- The statement that the present approach provides a fast alternative to FEA is not fully justified. In fact, as the authors mention in the discussion, many FE simulations are required anyhow for training, and their computational cost is presumably

alleviated by high-performance computing. But then, one may argue that standard FEA may be efficient enough to question the value of an ML approach for predicting only global stress-strain curves, particularly for relatively simple material behaviors (stable, small-strain elastoplastic materials). Perhaps multiscale analyses (e.g., FE2 methods) where the NN would replace the lower-scale simulations would serve as a better motivation.

- A five percent error for mesh convergence seems rather large for the present material response.

- The authors claim that "the sample scaling did not negatively affect the predictive capabilities of this model." However, the issue of scaling is not convincingly addressed. Fig. 1 provides some evidence of invariance for morphological parameters, herein presumed to fully encode the material response. However, there is no guarantee that the mechanical response will show the same convergence. It is known that non-linear, heterogeneous materials exhibit size effects, so the statement that the specimen can be printed at arbitrarily scaled-up sizes is far too strong. Indeed, for vanishing hardening responses, strains may localize, strongly challenging the notion of effective material properties and most of the presently drawn conclusions.

- In order to address the previous point, the authors may show invariance of FE simulations, assuming they remain close to experimental observations. However, the expected limitations in the presence of more complex responses (geometrical and/or material instabilities) should be clarified.

- Along these lines, the statements in the discussion concerning responses with damage and fracture are not justified; such cases would strongly complicate the analysis.

- As previously raised by Reviewer 3, there is no evidence that the four chosen scalars are particularly informative in comparison to the other 31. Is there any evidence that no other set of 4 parameters, picked from the remaining 31, would yield similar results?

- No evidence justifies the claim that even cases with anisotropy are accurately predicted.

- It is persistently claimed that the compressive strength of the material is being predicted. This terminology may be misleading, possibly alluding to failure strength, while the paper essentially focuses on predicting the stiffness and yield stress.

- The authors should consider including all ML predictions (beyond Fig. 6) in the supplementary materials.

- Finally, there are still several typos and some unclear sentences. A thorough proofreading process is required.

Reviewer #3

(Remarks to the Author)
The authors have improved their manuscript significantly in their revision to the level at which it is my pleasure to recommend it for publication.
The comments below are meant as recommendations, not requirements for publication for this reviewer.

One clarification comment related to Reviewer #3, comment #1 of the previous round:
p. 12, l. 291-297 of the current manuscript: In the reviewer's opinion it would be beneficial to try less than 4 Minkowski functionals. To the best of my understanding, the current hypothesis is "4 Minkowski functionals are sufficient for good accuracy of the ML model predictions in comparison to the full set of 35 parameters". But can this statement be strengthened? Are 4 Minkowski functionals also necessary? Is not ML capable of compensating for some effects/some of the MFs? Hence, if the statement could be strengthened to "4 Minkowski functionals are sufficient and necessary for a good accuracy of ML", I would find the main message of the entire manuscript significantly more convincing. I reckon, nevertheless, that this might incur some additional work on the side of the authors, without a guaranteed positive outcome.

A few textual details:
p. 4, l. 65: "functionals can fully effectively" --> "functionals can effectively"
p. 8, l. 158-161: the end of the sentence seems to be broken --> reformulate?
p. 9, l. 195: "velocity" --> "displacement"? J2 is rate independent, unless I missed something.
p. 9, l. 203: consider providing more details for the hardening law; current description is vague.

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The authors have improved the manuscript and addressed most concerns. However, some points require further revision or clarification, which, if properly addressed, should render the paper suitable for publication.

- It is not true that "a stress-strain curve provides an invaluable amount of information about its strength, deformation, toughness, fracture mechanisms, and much more." Stress-strain curves merely provide information about effective stiffness, strength, and toughness, but nothing can be inferred concerning fracture mechanisms. Indeed, all effective properties derived from stress-strain curves assume stable material behavior and are not representative in the post-critical stage. This point remains the main weakness of the paper since the proposed framework does not predict anything other than global compressive stress-strain curves. While this limitation is discussed by the authors, the value of the present approach remains quite vague.

- Strain localization is not a size effect per se; responses with strain localization exhibit size effects, such as strength or failure modes that depend on the ratio between the microstructural length scale and the structure size.

- While improved, the manuscript still contains grammatical errors. An honest and thorough proofreading process is required.

Reviewer #3

(Remarks to the Author)
The authors have sufficiently addressed all my comments.

Response to Reviewers: COMMSENG-24-0111-T

Dear reviewing committee,

We greatly appreciate the time, effort, and attention to detail you have all provided in reviewing this work. All of your comments were insightful and greatly helpful in focusing the scope of this work, and we feel that this manuscript has benefitted greatly from it. From the previous comments, a majority of our revisions have been structural in nature, with adjustments to title, abstract, introduction, and conclusion to better clarify the scope and impact of the research performed. Other adjustments have come in better explaining pipeline methodology and overall workflow, as well as individual clarifications on morphological calculations, machine learning, and finite element methodology. Again, we deeply appreciate the reviewers' comments, with specific addressing of reviewer points highlighted below.

Sincerely,

Corresponding co-authors

Dr. Winston Lindqwister Duke University Department of Civil and Environmental Engineering Winston.lindqwister@duke.edu	Mr. Jacob Peloquin Duke University Department of Mechanical Engineering jake.peloquin@duke.edu

Reviewer 1

#	Reviewer Notes	Author Response
1	Please double check the manuscript, some typos still remain.	Thank you for pointing this out, we've gone through the manuscript and corrected these. Our apologies!
2	The selection of the REV volume seems a bit arbitrary (pg. 7). For a ratio of 0.1 significant scattering can be seen in fig. 2. Please better justify the adopted choice.	Thank you for pointing this out—we indeed made an error in the submission of the original manuscript, the minimal volume fraction should be 0.2 instead of 0.1 (where the variance of the functionals existed below 0.01). We apologize for the confusion on this. To ensure this has not affected the overall results of the study, we checked the volume fraction of the scan subsamples and found that our samples selected were still within the acceptable REV range. The
3	It is not clear how the mesh is created. The creation of a mesh of a porous material from voxel data is not trivial and might lead to error if the features are very close. How do the Authors controlled the mesh generation? Is it a voxel-based mesh? Which finite elements have been adopted? Do the Authors noticed any locking effects due to plasticity?	We have clarified the meshing method on page 8, lines 170 - 177. Essentially, the voxels of the image stack are translated to hexahedron elements, as detailed in one of the cited works from the paper. Shear locking was minimal for the meshes created, though some samples displayed oscillations and discontinuous curves—these were removed from the dataset.
4	Many doubts remain also for the network training. Please explain better the procedure that has been followed, especially the stopping criterion used. Also, this type of problems are often non-convex and many sets of parameters might lead to low cost functions. Has this been taken into account? How do the Authors address this problem?	More information regarding the network training has been added in section II-E. Specifically, the network parameters that were varied during the grid search optimization are called out in Table II. The caption in Table II has also been updated to clarify how parameters were selected. Furthermore, the following information about model training is included in section II-E: - Network parameters that were varied - Cross validation - Training/Test split ratio - Pre-training (greedy) - Objective function used (Levenberg-Marquardt with MSE) - Stopping criterion for network training - Final selection of model based on performance and relative simplicity To address the non-convex nature of this type of problem, it is important to note that the authors used a cross-validated parameter grid search to select the highest performing hyperparameters of the network, while model training itself was performed with the Levenberg-Marquardt (LM) algorithm. Thus, selection of hyperparameters was based on a discrete 10x5x4 grid where continuous optimization was not performed. In the individual model training, the LM algorithm allows for initial values to be further from

		the global minimum by combining Gauss-Newton and gradient descent algorithms, and with added greedy layer-wise pre training of individual layers, the chances of approaching a global minima are improved.
5	In general, it is not completely clear what is the scope and the main goal of this work. As proposed, it allows to surrogate the determination of stress-strain pairs for a compressive test, however, for a comprehensive characterization of a material much more information are needed, such as tensile behavior, permeability, flexural behavior to name a few. Also, the choice of doing compressive tests on cubic samples is not optimal for, e.g., the Young's modulus due to the edge effects. Although the proposed work presents a methodology rather than an approach for a specific material, it is not clear what the breath of the work is. Do the Authors envision the adoption of this type of approach for a specific material? If not, what is the main aim of this method?	The reviewer is correct in understanding that this is a proposed general methodology rather than an approach for a specific material, and the authors apologize for the goal and scope of the work not being made clearer. The scope and main goal of the work have been more clearly defined in the abstract, introduction, and conclusion, as well as throughout the manuscript in certain areas. Particularly, the reviewer may find the following passage from the introduction helpful: “The overall goal of this research was to develop a systematic methodology for characterizing porous materials using scalar morphological descriptors and ML techniques to predict their compressive strength. By integrating physical and computational elasto-plastic experiments with advanced morphometric analysis, we aim to expedite the discovery of structure-property relationships in porous materials across a variety of fields including geomaterials, structural biomaterials, and energy storage materials. This approach seeks to overcome the challenges associated with the complexity and variability of porous media microstructures, enabling efficient screening of design spaces and enhancing predictive capabilities using succinct morphological descriptors and ML models.” Specific answers: - This work is not an attempt at comprehensive material characterization, but rather development of a method for comparing and predicting compressive properties of dissimilar porous materials. - The authors have clarified in the text that this methodology, through scalability and the universality of Minkowski functionals, could be applied to any arbitrary porous material as the prediction is based solely on geometry and base material properties. In the absence of physical samples, 3D printed surrogates were created to compare directly with FEA results. - Compressive tests on cubic samples are in fact quite common, and are often specified in ASTM standards for compressive testing of porous media (ASTM C109, ASTM D1621, ASTM C170, etc.)
6	The Authors also limit the study to elasto-plastic behavior, this can be a limiting assumption since damage might take place as well. Also, contact between different parts of the microstructure can take place.	Indeed, this study is limited to elasto-plastic behavior. We’ve updated the text to better reflect this, and hopefully this is more clear from the outset of the paper. Indeed, damage and contact elasto-plasticity may occur, although the stress-strain curves used in this study are not sufficiently deep in the plastic region where either of these phenomena are expected to constitute a significant response on the microstructure in question. However, we greatly appreciate you bringing up this point, and have added these limiting assumptions to the discussion portion of this manuscript**.**

7	Can the Authors comment on the expected precision of the CT scan images? This can limit the generality of the descriptors if some specific features are not captured.	The CT scan images used in this study all come from open repositories of porous media data, with these repositories linked to peer-reviewed studies at least attesting to the quality of the work performed. However, to the reviewer's point, differing scanning parameters can lead to many different scaling values for the studies in question, even if all studies that were used involve micro-computed tomography (which narrows the length-scale range of data). In order to maximize the generality of the dataset used, all scans were measured by pixel/voxel-based units of distance only, disregarding unit scaling from the scan metadata. One important facet of the measurements taken for the Minkowski Functionals is that all measurements are taken based on pixel-based calculation methods, speaking to the validity of treating all scan data as pixel-only.
8	Did the Authors try an inverse problem where given a stress-strain curve a set of descriptors is given? This might be an interesting extension to a-posteriori estimate the descriptors from experimental tests.	While the inverse problem is beyond the scope of this particular manuscript, it is a natural next step of this work and one we are very interested in pursuing for future work. One of the primary challenges would be creating a synthetic microstructure generator that is sensitive to Minkowski functionals only—these functionals, while mathematically independent, are difficult to isolate in image generation functionality. This is mentioned in future works, and we appreciate the reviewer taking the time to consider where this work has potential to grow.

Reviewer 2

#	Reviewer Notes	Author Response
1	The novelty is not sufficiently clear. What is the benefit of this approach vs. other (rather abundant) works in the literature that also rely on ML to predict global stress-strain responses?	We appreciate the reviewer's comment and recognize that the main message of this work isn't completely clear. Hopefully, the updates to the manuscript's introduction, abstract, and overall structure will make these features a bit clearer. Specifically, to this reviewer's comments, we would like to clarify that the novelty of these findings is the fact that we could find strong predictive power in describing the compressive strength of porous materials via four scalar descriptors of geometry. In short, the combination of these descriptors and a machine learning model creates a fast, relatively accurate means for recreating the mechanical response of a porous material based solely on its geometry. Because these are scalar values, they are significantly more data-portable than full sized images, and from the results of this study, work well even in cases of relatively large 3D scans. Naturally, this implies that the next result would be an analytic approach to linking compressive strength to Minkowski functionals (MFs). However, one has not been found at this point in time, and we feel that these results of using MFs and other scalar descriptors serves as a worthy enough advancement in porous media analysis to report to the scientific community.
2	I struggle to see the added	The authors appreciate the reviewer's consideration of novelty with

	value of an ML approach for reproducing the global stress-strain behavior. In particular, if the only goal is to capture the effective stress-strain behavior, one may use simple empirical models or constitutive plasticity theory at the material point level to yield the target response without resorting to FEA. Of course, for more ambitious objectives, e.g., predicting full-field quantities or averaged properties that require full-field microstructural solutions, one could argue that FEA simulations are expensive and that ML approaches may come in handy, but this does not seem to be the case here. What is the added value of the present ML approach in this context?	respect to ML approach. We invite the review to briefly read the reviewer 1, #5 note above to help clarify the scope of this work. As for the ML itself, the power of the developed model lies with the speed and simplicity that it offers. Using only the 4 Minkowski functionals to predict compressive behavior, and with predictions being made in milliseconds, large amounts of dissimilar porous media data can be screened or compared. Future applications of such a framework may include fast analysis and weakness detection in large porous media samples through sub-sampling or introduction into material-based large language models (LLMs) such as MechGPT.
3	The description of the geometric parameters is a little vague. I suggest adding a few details. For instance, how is the mean curvature captured by the grain size? How is the mean breadth computed if not all solid domains are convex?	We've added some clarification for the descriptors on page 7, as well as more clear connection between mean curvature and grain size. As for the breadth computation, grains are calculated via a watershedding algorithm (noted on page 7 line 142). This algorithm takes a nearest distance-to-boundary calculation for delineating what is considered a grain, essentially cutting a non-convex shape into multiple convex shapes or grains. Any shape made up of a sum of convex bodies is considered polyconvex, which still falls within the domain of Hadwiger's theorem.
4	Why is the J2 plasticity model chosen? Again, this point may reflect the fact that only the global response is studied, which can be captured in many ways.	The J2 plasticity model was selected because these materials were modeled as connected bodies. Essentially, while their original scans came from a wide variety of materials, we wanted to look specifically at 3D printed representations of these materials to isolate geometry in our model's fitting. Thus, the J2 plasticity criterion was selected. We agree that there are many ways to study the global response of materials under compression, and there is indeed a wealth of literature available on this topic. However, we felt that our work stood on its own in examining specifically Minkowski functionals and their relationship to elasto-plastic uniaxial compression.
5	Simulations with finer meshes could be conducted to confirm that the results of poorly predicted microstructures are due to over-coarsening.	Indeed, we performed simulations with finer meshes to examine for over-coarsening. While this did help somewhat with divergent scans, ultimately the samples that proved the most troubling were those with highly convex geometries, as detailed on pages 14 - 15.
6	Can the authors comment on the dependence on curing time	We appreciate the reviewer's question regarding dependence of these parameters. This analysis was included after receiving

	and temperature, particularly considering that only morphological parameters are inputs?	feedback during a previous review process with a different journal. After reading the feedback on our most recent submission, we have decided to remove this portion of the study from the manuscript as we believe, much like **Reviewer 3 in their #2 note below**, that it clouds the scope and goal of the paper. For this manuscript, printing parameters like curing time and temperature were held constant when comparing 3D printed materials to the computational results. However, including printing parameters like these would indeed be a worthwhile study for future models trained on primarily physical 3D print data. Unfortunately, this is beyond the scope of this manuscript.
7	At the beginning of page 8: did the authors mean "exponentially [decrease] computational time"?	Yes, the authors meant to state that the reduced order of nodes would decrease computational time. We apologize for this mistake in the original submission and have corrected the revised version. Thank you!

Reviewer 3

#	Reviewer Notes	Author Response
1	It is not clear from the manuscript whether performance of multiple feature combinations have been tested in place of microstructural descriptors. The authors mention on page 10 that “multi-dimensional search grid of 200 parameter combinations was developed”, but later comment that “the model was retrained with only the four Minkowski functionals”, which makes it somewhat inconclusive. If other feature combinations have not been tested, then the main statement of the manuscript is in my opinion not substantiated, since the authors do not convincingly show/demonstrate performance of other feature combinations (e.g., three instead of four Minkowski functionals, or other combinations of geometrical descriptors or completely different descriptors, cf. comment 20 below). This would then weaken the main hypothesis of the manuscript and would make the entire study fairly inconclusive.	We apologize that the model training and parameter selection was not made clearer in the original manuscript. Additional information regarding model training has been added to the manuscript in Section 2E to clarify how the grid search was conducted and what input features were used. Specifically, a total of 35 microstructural features were extracted for each sub-sample. The “200 parameter combinations” consisted of three parameters for the neural network: number of nodes/layer, number of layers, and type of layers. During model training, initially all 35 microstructural features were used as inputs, and the 200 neural network combinations were all trained and then evaluated for performance with these 35 inputs. Then, the process was repeated using just the 4 Minkowski functionals as feature inputs, but again with all 200 neural network combinations. The best performing combination of network parameters was selected from each training process, and these results were presented in Figure 5 and discussed. The minimal reduction in performance between using all 35 features and the 4 Minkowski functionals highlights the power that these functionals have in predicting compressive performance, and that has been clarified as the main goal/scope of this work.

2	The role of experiments is not clear. The authors seem to employ experiments to verify numerical predictions against experimental measurements (or, vice versa), and to identify material parameters of the bulk material model. This seems to be unnecessary since correlation between geometrical features and stress-strain relationships through NN is being tested and any ad-hoc material (with plasticity) would serve the purpose equally well. The study and the main results do not seem to benefit from the conducted experiments, and even seem to cloud the main message of the submitted manuscript.	The authors agree that the form and structure of the initial manuscript clouded the goal and scope of the work, and we appreciate the reviewer pointing this out to improve the manuscript presentation. The role of experiments has been clarified in paragraph 2 of Section 2D, and we believe that the addition of physical experiment using 3D printed representations of the microstructures to validate and compare to FEA simulation provides a valuable example of how scalable and applicable our developed model is, particularly in the absence of the physical raw material samples that CT scans were created from.
3	The structure of the paper is not very clear and resembles a report rather than a manuscript with a carefully build argumentation; see also comment 15 below	The authors have made substantial revision to the manuscript, removing sections regarding “calibration curves” and attempting to improve the flow of the presentation. The experimental methodology has been moved to the end of the materials/methods section in a separate subsection. Words like “project” have been modified to highlight the basic scientific nature of this study, though we believe that, as mentioned by Reviewer 1 in their #5 note above, it is important to clearly state the goal and scope of the work.
4	Page 4: “Minkovski functionals can fully describe a microstructure”. Certainly, a too strong statement, in addition contradictory with latter statement on page 12 “…it was hypothesized that these four functionals…”. Please clarify and correct.	The authors acknowledge that “fully describe” may indeed be too strong a statement, particularly if description of a microstructure is taken to mean more than just geometry. As such, we have adjusted this sentence to read: “d + 1 Minkowski functionals can effectively characterize the geometric shape and composition of a microstructure”, which is a claim supported by mathematical theory and Hadwiger’s theorem itself.
5	Introduction: it would be good to clearly specify the considered setup early on in the manuscript and mention that only uniaxial tests are considered, leading to only one component of stress-strain relationships.	This is a great point. We have clarified in multiple places including the abstract, introduction, and methods that strictly uniaxial compressive loading was analyzed in this study. Additionally, we have clarified within the title that compressive loading was analyzed.
6	Section B: representative elementary volume. How does this relate to typically used representative volume element (RVE), known from computational homogenization?	The terms “representative volume element” (RVE) and “representative elementary volume” (REV) are used interchangeably in the field of material science.

7	Section C: "…the image was binarized to separate the solid…"; how exactly has this been done, what was the threshold value, etc.? Similar details may significantly influence the results. Additional question: could the authors elaborate on how the proposed methodology would work for multi-material systems? Binary composites considered, i.e., bulk/void, seem to be fairly limited for practical applications.	We agree with this point that image processing parameters would have a significant impact on the 3D scans that were generated. From our comment with Reviewer #1, note 7, the individual parameters are available on the cited repositories, collected in the Appendix of the work (Appendix A). To the scope of this work, we are solely using the CT scan dataset as a dataset of real-world geometries to populate our study. Because we are 3D printing these scans, the structures tested physically are based on the scan parameters chosen, and thus the threshold values don't have much of an effect in the context of our specific study (as in, a study that solely considers geometry, which is relevant for 3D printed applications). For multi-material systems, grayscale thresholds in the binarizing step of image analysis would allow for the differentiation of different material types, which would be separated easily via MOOSE's image reader. Additional features to consider in a model would be the idealized stiffness of the two (or more) composite materials in question, as well as what percentage these materials make-up of the bulk material. Naturally, a study that includes simulations for varying stiffness would be an important next step for this work, and it is one we are interested in pursuing in future efforts. However, it is beyond the scope of a study focused mainly on porous microstructure geometry.
8	Page 7: "motion-invariant—or isotropic—microstructure". Could the authors specify more what they mean with isotropic? There seems to be a collision in terminology (the text seems to discuss group of transformations, while isotropic microstructures can be interpreted also as those that result in isotropic effective material properties/behaviour).	Thank you for pointing this out. We refer to isotropic as a geometric measurement that does not vary significantly in any direction. For example, when measuring porosity, if we took 2D measurements of porosity of the sample on random x, y, and z planes, the porosity measurement should end up being roughly the same. While this is true in our large samples, we were worried about the sub samples potentially losing overall anisotropy, as their smaller volume would be more sensitive to stochastic variations. Some of the other features included in the study are directionally dependent (i.e. box x min), so we felt that including these may offer some insight on anisotropic effects.
9	Page 7/8: "To speed up the collection … to potentially reduce the order of nodes on the mesh and to exponentially increase computational time …" à decrease computational time?	The reviewer is correct, this should have read "decrease computational time". This point has been similarly made by Reviewer 2 in their #7 note above and it has been corrected in the manuscript.
10	Page 8: "finite element results were compared to physical experiments": the role of experiments is not clear in the manuscript, cf. point 1. Please clarify.	**Please see response to Reviewer 3, note #2 above.**
11	Table I: entries in the last two rows and first three columns seem to be missing.	Clarification has been added in the Table caption and within the table by filling in blank spaces with "-" to indicate that these are not missing cells, but rather that the table contains a list of three different parameters and their values.

12	Table II: again, entries missing and not clear how to interpret the blank space in the table.	The original Table II has been removed from the manuscript, **see response to Reviewer 3, note #2 above.**
13	Section B: “experimental validation of FEA” not clear why this section is present and why FEA needs to be validated.	**Please see response to Reviewer 3, note #2 above.**
14	Page 12: “scaling the size of the printed cube had a minimal effect”. Is this assuming fixed geometry? If yes, then this exercise confirms scale-independence of the employed constitutive law, which is clearly to be expected a priori. The follow-up statement “This supports the assumption … is representative” seems to be then wrong. Please clarify and elaborate.	**Please see response to Reviewer 1, note #7 above.**
15	Page 13: “the overall goal for the project”, “The results of this project” seem to suggest a report-like approach towards the submitted manuscript; please refer to comment 3 above and correct/clarify.	**Please see response to Reviewer 3, note #3 above.**
16	Page 13: “demonstrate not only the capabilities of Minkowski functionals…”. This statement does not seem to be substantiated properly, since performance of Minkowski functionals has not been compared against other descriptors. It can be hypothesized that deep NN could compensate for the effect of using other descriptors. See also comment 1 above.	**Please see response to Reviewer 3, note #1 above.**
17	Page 14: “The minimal reduction in performance of the model when using only the four Minkowski functionals demonstrates…”; cf. comment 16. The claim is not well substantiated.	**Please see response to Reviewer 3, note #1 above.**
18	Page 15: “the focus of this work was to link theory, simulation,	**Please see response to Reviewer 3, note #2 above.**

	experiment"; the role of either simulation or experiment is not clear. Why both and not, for instance, numerical simulations only?	
19	Conclusions: could the authors comment on the accuracy of predictions in terms of different loading conditions (assume, for instance, a typical RVE element in computational homogenization) and not a single tensile test, employing all deformation modes. To which extend are achieved conclusions transferrable to such situations?	We appreciate the insight of this comment, as it is a natural evolution of the work presented—future directions for this work would be examining the predictability of Minkowski functionals for the entire yield envelope of a material, which essentially accounts for a full triaxial loading battery. In theory, Minkowski functionals should work just as well for a yield envelope case, as the requirement for the material to be isotropic should lead to comparable results across multiple loading conditions. This would be an exciting study that we hope to pursue in the future!
20	Appendix B: It is not clear to the reviewer what was the motivation for including the listed descriptors. Why not including standard descriptors corresponding to random microstructures (see, e.g., Torquato, Random heterogeneous materials, for more details).	The descriptors used for this study were selected as descriptors commonly used in porous media and image analysis, derived from the works of *Burger et al. (2008)*, *Lang et al. (2001)*, *Serra (1982)*, and the extensive works of *Legland et al.* In morphometric analysis of porous media (works cited in the manuscript and through MorphoLibJ). We feel that this set of descriptors offer a serviceable complement to Minkowski functionals and are all readily calculated from stacks of binarized images. Ultimately, while Minkowski functionals offer bulk descriptions of the polyconvex space, the additional descriptors are essentially averages of grain-defined properties through the porous media, and in our opinion are sufficient for the scope of this work. However, we do appreciate the suggestion of examining additional microstructural features. Descriptors such as lineal path function and average chord length would fit nicely with future studies on applications of anisotropic materials, and we have mentioned as such in the conclusion.
21	p. 12: "the reactive force" → "the reaction force"; please correct elsewhere.	These corrections have been made throughout the manuscript, thank you!

Response to Reviewers: COMMSENG-24-0111-T

Dear reviewing committee,

We are again deeply appreciative of the constructive feedback for this manuscript. Thanks to your efforts, we feel that this work has improved greatly in terms of clarity and concision. Furthermore, the additional suggestions from the reviewers have helped pull the full story of this work together. Attached below are our responses to individual comments from the reviewing committee. We hope the changes in this document reflect the referee comments and look forward to hearing any and all future feedback on this work.

Sincerely,

Corresponding co-authors

Dr. Winston Lindqwister Duke University Department of Civil and Environmental Engineering Winston.lindqwister@duke.edu	Mr. Jacob Peloquin Duke University Department of Mechanical Engineering jake.peloquin@duke.edu

Reviewer 1

#	Reviewer Notes	Author Response
1	*Most of the remarks raised have been addressed, however it is still not clear how this approach should be able to improve discovery of new materials (reply to point 5 raised) only considering the compressive behavior. There is no evidence that this is enough, considering that the method does not take into account any other mechanical behavior or characteristic than the compressive stress-strain curve. The Reviewer suggests to reduce the breath of the method to those application where only the compressive behavior characterized by J2 plasticity is needed/relevant.*	We appreciate the feedback from Reviewer #1 regarding the breadth of this work. We have made changes throughout the manuscript to highlight the focus and applicability of this work. Specifically, we have focused on how the ML predictions can be utilized in scenarios where many different material samples are being investigated and/or where predictions are required at speeds much faster than FEA or experimental methods. We have also highlighted that this framework made for compressive stress-strain behavior provides a foundation, and that future models that predict other mechanical behavior, or perhaps a single model that predict multiaxial stress states, would be the logical next step: - For specific instances of changes addressing the applicability of the proposed framework, please see response to Reviewer 2, comment 1 below.

Reviewer 2

#	Reviewer Notes	Author Response
1	*The statement that the present approach provides a fast alternative to FEA is not fully justified. In fact, as the authors mention in the discussion, many FE simulations are required anyhow for training, and their computational cost is presumably alleviated by high-performance computing. But then, one may argue that standard FEA may be efficient enough to question the value of an ML approach for predicting only global stress-strain curves, particularly for relatively simple material behaviors (stable, small-strain elastoplastic materials). Perhaps*	Utilization of FE2 methods with NN serving as the lower-scale simulations is a very interesting idea. We would like to address this reviewer comment in two ways: 1. This proposed methodology sounds very promising, particularly if small-scale simulations at the microscale are replaced with a NN (after sufficient training of course), while larger scale simulations such as meso to macroscale that account for more complex material behavior or perhaps multimaterial interaction. The authors greatly value this suggestion from the reviewer, and have made attempts to highlight the focus and potential application of this work in such a manner in several spots throughout the manuscript: - Lines 24-27, Abstract, “This framework…” - Line 44, Section I, “…in certain applications.” - Lines 91-92, Section I, “…with the intent…” - Lines 95-96, Section I, “As a starting…” - Line 98, Section I, “…calculation…” - Lines 100-107, Section I, “This approach seeks…” - Line 200, Section IID, “…in similar form…” - Lines 361-370, Section IV, “The approach taken…” - Lines 473-477, Conclusion, “Nonetheless,…” 2. Perhaps the utility of this framework is best thought about when considering number of samples as a variable. While we did not originally include such an analysis, we have attempted to highlight the preferability of a NN over FEA for instances where many future

	multiscale analyses (e.g., FE2 methods) where the NN would replace the lower-scale simulations would serve as a better motivation.	simulations are likely to be required (perhaps such as the FE2 scenario mentioned by the reviewer): - Figure 7 - Lines 340-345, Section IV, “Utilization of high performance…” - Lines 346-360, Section IV, “In Figure 7A,…”
2	*A five percent error for mesh convergence seems rather large for the present material response.*	While we agree that a five percent error is on the larger end for the presented material response, we feel from the mesh convergence data (provided in the supplemental material) that this response is still adequately accurate, given the dramatic gains in simulation time. Furthermore, this same coarsening was applied to all samples, so any error found with an over-coarsened mesh should not affect the ultimate experimental results. Furthermore, from the results displayed in Figs. 4 and 6, the ultimate computational results showed strong agreement with the values derived from the 3D printed experimental verification set.
3	*The authors claim that "the sample scaling did not negatively affect the predictive capabilities of this model." However, the issue of scaling is not convincingly addressed. Fig. 1 provides some evidence of invariance for morphological parameters, herein presumed to fully encode the material response. However, there is no guarantee that the mechanical response will show the same convergence. It is known that non-linear, heterogeneous materials exhibit size effects, so the statement that the specimen can be printed at arbitrarily scaled-up sizes is far too strong. Indeed, for vanishing hardening responses, strains may localize, strongly challenging the notion of effective material properties and most of the presently drawn conclusions.*	See response to Reviewer 2, comment #4 below.
4	*In order to address the previous point, the authors may show invariance of FE*	As referenced by the reviewer, demonstrating scale-invariance of the morphological parameters was the primary goal of the REV analysis. Scale invariance of mechanical performance is indeed a much more complicated claim to prove, and the authors appreciate the reviewer for pointing this out.

	simulations, assuming they remain close to experimental observations. However, the expected limitations in the presence of more complex responses (geometrical and/or material instabilities) should be clarified.	Our goal is to demonstrate that results obtained from experimental evaluation of the 3D printed analogues, the FEA simulation, and the morphology-informed ML predictions form a three-way validation between theory, computation, and experiment. We have made changes to the manuscript to highlight this, and to not give the sense that we believe this framework could be used at any length scale for any material created using any processing method: - Line 328, Section IIIC, "…through scaling…" - Lines 371-375, Section IV, "The results of…" - Lines 381-390, Section IV, "While the REV…" - Lines 469-473, Conclusion, "While the model…"
5	*Along these lines, the statements in the discussion concerning responses with damage and fracture are not justified; such cases would strongly complicate the analysis.*	Perhaps the previous form of this statement was misleading. The authors intended to highlight the fact that, in general, high-performance computing clusters and utilization of parallel core multithreading and multiprocessing would reduce simulation time for the computational finite element analysis. To rephrase: "This (utilization of high-performance computing clusters) would be especially important in simulations that take into account other mechanical phenomenon beyond elasto-plastic behavior, such as damage, fracture, and internal contact". This phrase was intended to claim that for even more complicated analysis like damage, fracture, and contact, these clusters would be crucial in order to generate the required data for training a model in a timely manner. Simply put: for more complicated analysis, more compute power would be needed to generate the same amount of compressive stress-strain data. We have modified this section of the manuscript to more accurately reflect this claim: - Lines 322-326, Section IV, "Utilization of high performance…"
6	*As previously raised by Reviewer 3, there is no evidence that the four chosen scalars are particularly informative in comparison to the other 31. Is there any evidence that no other set of 4 parameters, picked from the remaining 31, would yield similar results?*	We have performed additional analysis to highlight the relative performance of the Minkowski functionals, and have discussed in the manuscript instances where more than these four input features may be desirable: - Figure 5 - Lines 173-174, Section IIC, "…for ML model training…" - Lines 298-308, Section IIIB, "To evaluate the…" - Lines 391-417, Section IV, "The performance of the ML…" - Lines 463-477, Conclusion, "This study has…"
7	*No evidence justifies the claim that even cases with anisotropy are accurately predicted.*	Please see response to Reviewer 2, comment 6 above, this claim has been removed and readdressed.
8	*It is persistently claimed that the compressive strength of the material is being predicted. This terminology may be misleading, possibly alluding to failure strength, while the paper essentially focuses on predicting the stiffness*	We acknowledge that author's comment and have made further indications throughout the manuscript that what was predicted in this study was *compressive stress-strain behavior of elasto-plastic porous materials*, *specifically through the linear regime and onset of nonlinear plastic regime.* Furthermore, we have removed instances where it was claimed that "strength" or "performance" was predicted. Examples of these instances are listed below: - Lines 1-2, Title - Lines 22-27, Abstract, "…generating material stress-strain…" - Lines 91-92, Section I, "…with the intent…"

	and yield stress.	- Line 95, Section I, "…compressive loading" - Lines 185-186, Section IIC, "…generate a…" - Line 198, Section IID, "…compressive stress-strain behavior…" - Lines 310-313, Section IIIB - Lines 463-467, Conclusion, "This study has demonstrated…"
9	*The authors should consider including all ML predictions (beyond Fig. 6) in the supplementary materials.*	We have included predicted vs. simulated stress-strain curve plots for all 54 evaluation set materials in Appendix C, and referenced in the manuscript: - Lines 300-301, Section IIB, "Predicted vs. simulated…"
10	*Finally, there are still several typos and some unclear sentences. A thorough proofreading process is required.*	Thank you for pointing these out, we have taken time to further proofread the manuscript to be of an acceptable quality for submission.

Reviewer 3

#	Reviewer Notes	Author Response
1	*One clarification comment related to Reviewer #3, comment #1 of the previous round:* *p. 12, l. 291-297 of the current manuscript: In the reviewer's opinion it would be beneficial to try less than 4 Minkowski functionals. To the best of my understanding, the current hypothesis is "4 Minkowski functionals are sufficient for good accuracy of the ML model predictions in comparison to the full set of 35 parameters". But can this statement be strengthened? Are 4 Minkowski functionals also necessary? Is not ML capable of compensating for some effects/some of the MFs? Hence, if the statement could be strengthened to "4 Minkowski functionals are sufficient and necessary for a good accuracy of ML", I would find the main message of the entire manuscript significantly more convincing. I reckon, nevertheless, that this might incur some additional work on the side of the authors, without*	Please see response to Reviewer 2, comment 6 above.

	a guaranteed positive outcome.	
2	*A few textual details:* *p. 4, l. 65: "functionals can fully effectively" --> "functionals can effectively"* *p. 8, l. 158-161: the end of the sentence seems to be broken --> reformulate?* *p. 9, l. 195: "velocity" --> "displacement"? J2 is rate independent, unless I missed something.* *p. 9, l. 203: consider providing more details for the hardening law; current description is vague.*	Thank you for pointing these out, these comments have been addressed.

Dear editorial committee,

We deeply appreciate the time and effort that has been put in to evaluate this work. All feedback has been invaluable in honing the message of this manuscript and allowing for a more complete picture of our research to shine through. Below are our responses to the concerns raised by the reviewers:

- It is not true that "a stress-strain curve provides an invaluable amount of information about its strength, deformation, toughness, fracture mechanisms, and much more." Stress-strain curves merely provide information about effective stiffness, strength, and toughness, but nothing can be inferred concerning fracture mechanisms. Indeed, all effective properties derived from stress-strain curves assume stable material behavior and are not representative in the post-critical stage. This point remains the main weakness of the paper since the proposed framework does not predict anything other than global compressive stress-strain curves. While this limitation is discussed by the authors, the value of the present approach remains quite vague.

This is absolutely true, and we feel the mention of fracture mechanisms confuses the main findings of this research. As such, we have decided to remove mention of fracture mechanisms (pg. 5 line 86) and instead focus on the many important properties of material stress-strain curves, namely its indicators of strength, deformation, and rigidity.

To the reviewer's second point:

- Strain localization is not a size effect per se; responses with strain localization exhibit size effects, such as strength or failure modes that depend on the ratio between the microstructural length scale and the structure size.

We agree with this clarification, and hope the additions on page 14 lines 368-377 and adjustments in the conclusion paragraph (pages 16 and 17, lines 423 – 443) recontextualize our statements regarding size effects and strain localization.

In the context of this work, this is a first study offering some responses for the simple case of uniaxial compression of porous media. The extension to generic three-dimensional loading conditions and variable failure modes (like shear localization and fracture) requires generalization of the framework to account for more elaborate models including information on pressure sensitivity[1] and how the microstructural morphometers evolve to capture ultimate shear band thickness[2], which is indeed planned for future contributions. Nonetheless, the fact that the current study can correlate certain aspects of the microstructure with the Uniaxial Compressive Strength (UCS) is a very big step forward,

especially since the UCS is considered to be a key design factor for many engineering applications.

Finally, we have done another pass through the manuscript to correct any grammatical or spelling errors. We hope we have adequately addressed these.

As always, we greatly appreciate the feedback on this work. We thank the editors and reviewers for their time and efforts in ensuring the scientific rigor of this manuscript.

Sincerely,

Winston Lindqwister
Duke University
Department of Civil and Environmental Engineering

Jacob Peloquin
Duke University
Department of Mechanical Engineering and Material Science

References

[1] Sari, Mustafa, Sotiris Alevizos, Thomas Poulet, Jack Lin, Manolis Veveakis. 2020. “A Visco-Plastic Framework for Interface Processes in Sedimentary Reservoir Rocks at HPHT Conditions.” Geomechanics for Energy and the Environment 22 (May): 100165.

[2] Jacquey, Antoine B., Hadrien Rattez, and Manolis Veveakis. 2021. “Strain Localization Regularization and Patterns Formation in Rate-Dependent Plastic Materials with Multiphysics Coupling.” Journal of the Mechanics and Physics of Solids 152 (104422): 104422.