Prediction of microstructural representativity from a single image

Amir Dahari[1], Ronan Docherty[1,2], Steve Kench[1], Samuel J. Cooper[1]

[1]Dyson School of Design Engineering, Imperial College London, London SW7 2DB, UK
[2]Department of Materials, Imperial College London, London SW7 2DB, UK
a.dahari@imperial.ac.uk

In this study, we present a method for predicting the representativity of the phase fraction observed in a single image (2D or 3D) of a material. Traditional approaches often require large datasets and extensive statistical analysis to estimate the Integral Range, a key factor in determining the variance of microstructural properties. Our method uses a recently developed AI generated image library MicroLib [2] to validate our approach which leverages the Two-Point Correlation function to directly estimate the variance from a single image (2D or 3D). The AI-generated library of realistic microstructures facilitates the creation of materials of arbitrarily large size, thereby supporting representativity analysis and validation. This new paradigm allows us to trust our method for phase fraction prediction from a single image with associated confidence levels. We further test our method using open-source datasets, both experimental and simulated, demonstrating its efficacy across diverse microstructures. The problem of undersampling and our proposed solution can be seen in Figure 1.

Microscopy is an important aspect of material's characterisation, allowing the statistical analysis and modelling of a material microstructure. For example, performing X-ray computed tomography (XCT) on battery electrodes can generate datasets which help us to understand processing-structure relationships, and enable electrochemical modelling to better understand performance limitations. Similarly, scanning electron microscopy (SEM) of steels has been used to understand phase distributions and model relative permeability. The characterisation-modelling-design workflow is ubiquitous in the field of material science.

The confidence analysis of metrics extracted from micrographs or microstructures is largely lacking in literature. The time and resources required to collect the many data repeats needed to determine confidence intervals poses an often-prohibitive barrier. Thus, it may sometimes be undesirable to discover that your technique requires an order of magnitude more data to give reasonable confidence in the results. To this end, we often choose to simply bury our heads in the sand, and report results with

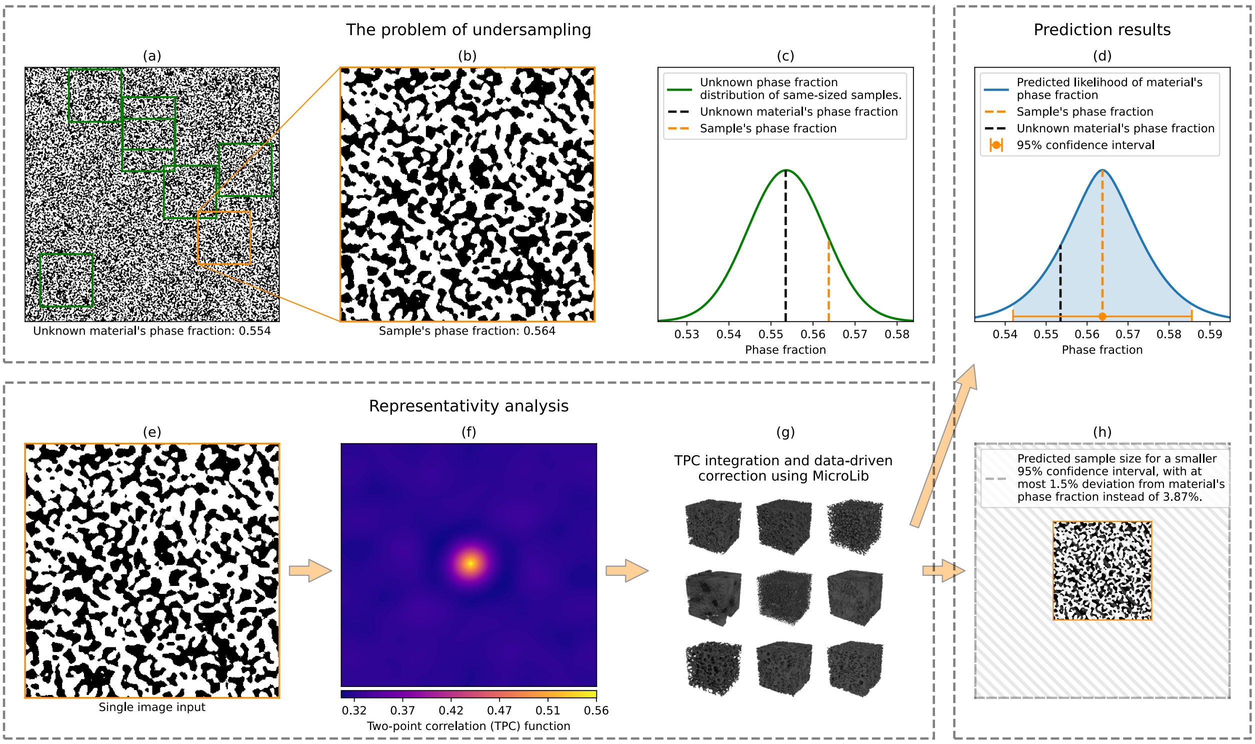

Figure 1. An outline of the undersampling problem and the solution presented in this paper. The difference between the sample and bulk material phase fractions is unknown, necessitating confidence bounds on their proximity, to infer conclusions about the material (The problem of undersampling). By utilizing both the TPC and a data-driven analysis of MicroLib [2] in a statistical framework (The representativity analysis), both the confidence bounds to the material's phase fraction and the image size needed for a specified deviation from the material's phase fraction are determined (Prediction results).

no thought given to sample representativity.

This can have dire consequences, as even small changes in phase fraction can significantly impact a material's properties and therefore performance. For example, in the case of batteries, phase fraction of active material directly determines the theoretical capacity of an electrode. Stainless steel's corrosion resistance or magnetic material's strength and behavior are all sensitive to tiny phase fraction changes. Precise control over phase fractions is therefore essential for optimizing materials for specific applications, ensuring optimal performance and reliability. To trust these performance results, we must understand the statistics of our microstructural measurements.

In this paper we propose ImageRep [1], for estimating confidence in phase fraction from a single micrograph or microstructure (2D or 3D). Our approach was developed using a database of microstructures, MicroLib [2] that was created using SliceGAN [4] generators. Each entry has a trained Generative Adversarial Network (GAN) that allows the synthesis of samples of any size. We leverage this capability to construct a model that directly relates the two-point correlation function (TPC) of an image to the confidence level in representativity of the phase fraction observed in that image. To guide practitioners, our method also predicts the image size required to obtain a smaller uncertainty in phase fraction, which can direct precise data collection campaigns after a small initial data collection study.

The single image requirement significantly reduces the data requirements for representativity analysis, providing a practical tool for material scientists and engineers working with limited microstructural data. The validation process to our method, tested over thousands of experimental and simulated materials, is presented in Figure 2.

In the age of high-throughput materials design using generative AI [3], incorporating a fast representativity prediction into the design cycle can provide guidance on the optimal size of generated materials, enhancing the efficiency of the process.

Pivotal to this effort, we have created a web-application, www.imagerep.io, for quick, simple and informative use of the method.

References

[1] Dahari, A., Docherty, R., Kench, S., & Cooper, S. J. (2024). Prediction of microstructural representativity from a single image. *arXiv preprint* arXiv:2410.19568

[2] Kench, S., Squires, I., Dahari, A., & Cooper, S. J. (2022). MicroLib: A library of 3D microstructures generated from 2D micrographs using SliceGAN. *Scientific Data* 9 (1), 645

[3] Kench, S., Squires, I., Dahari, A., Planella, F. B., Roberts, S. A., & Cooper, S. J. (2024). Li-ion battery design through microstructural optimization using generative AI. *Matter*, 7(12), 4260-4269.

[4] Kench, S., & Cooper, S. J. (2021). Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion. *Nature Machine Intelligence*, *3*(4), 299-305.

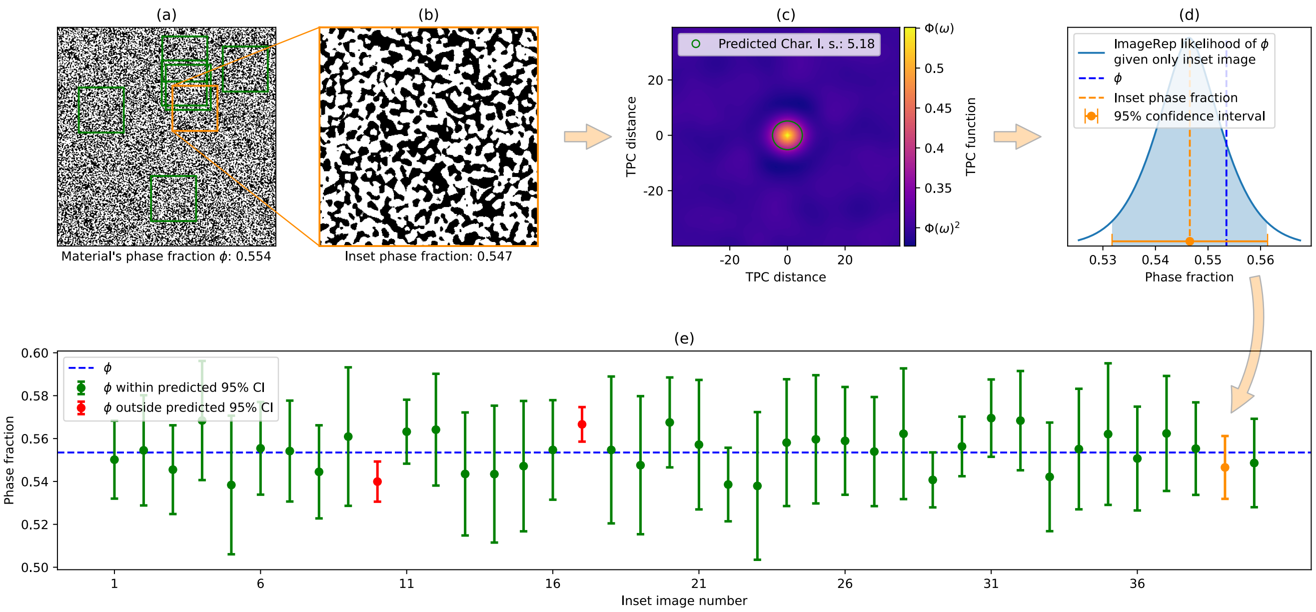

Figure 2. Overview of the ImageRep method validation process: For a given confidence level such as 95%, the method needs to produce correct results 95% of the time. The proposed method calculates the TPC function (c) and arrives at the predicted confidence interval (d). The bulk material's phase fraction should fall within the predicted confidence interval in 95% of cases. As a toy example, (e) presents 40 different confidence intervals for 40 different random image samples taken from (a), showing that in 38/40 = 95% of the cases, the material's phase fraction is within these intervals, as desired. the corresponding interval for image (b) is presented in inset image number 39.