MORPHOLOGICAL DEEP LEARNING MODELS FOR MICROSTRUCTURE CHARACTERIZATION AND GENERATION

MODÈLES D'APPRENTISSAGE PROFOND MORPHOLOGIQUES POUR LA CARACTÉRISATION ET LA GÉNÉRATION DE MICROSTRUCTURES

***Composition de l'équipe d'encadrement / supervising team**:* Jesus Angulo (CMA, directeur de thèse), François Willot (CMM, co-directeur), Santiago Velasco-Forero (CMM), Francesco Delloro (MAT)

Nom du laboratoire et équipe d'accueil / laboratory : Centre de Mathématiques Appliquées (CMA)

Sujet de thèse / PhD thesis

Contexte / Context

The modeling of heterogeneous materials such as composites, porous structures or granular media, and their mechanical properties, typically involves different length scales. At the macrocopic scale, a material is often treated as a continuum with uniform properties, allowing for simplified numerical computations. At the mescoscale, the material may experience strongly inhomogeneous local stress and strain. Examples include inelastic deformation through the motion of defects, brittle or ductile fracture when damage starts to localize, and crack propagation or nucleation. All these mechanisms are driven by the local stress state, which is itself a complex result of the load distribution within the microstructure of the material. Modern material processing methods aim at tailoring microstructure to achieve desired mechanical properties, which are ultimately used in reverse-design or to control and assess the material's mechanical response. Yet, industrial (and natural) materials generally present *highly complex* micro-geometries; characterizing and simulating them.

Traditionally, this task has been performed by viewing the microstructure as a realization of a *random set* in a "representative volume element" and involves the following step: (i) the computation of a set of geometrical indicators (granulometry, covariances) that provide a partial characterization of the micro-geometry; this is typically obtained from experimental images, using image analysis techniques such as segmentation and morphological operators; (ii) choice of a mathematically well-defined random sets [1,3] involving a small set of parameters as a model for the microstructure; (iii) numerical optimization of the model parameters to recover the observed characteristics; (iv) validation of the model with respect to experimental images/tests. Deep learning approaches, instead, such as autoencoders, learn features from images, represented in a latent space, that can be used to generate new images that resemble an experimental or synthetic dataset. Generative adversarial networks (GANs) [9] or more recently diffusion models, are other ways to generate complex microstructures with the same statistics as the training set. These models do not involve any arbitrary choice on the underlying statistics that are relevant to describe microstructures and, since they involve many degrees of freedom, they are able to capture complex morphology. However, because of their black-box nature, it is difficult to identify which spatial features must be satisfied to generate consistent microstructures, that can be understood by a human. This problem is sometimes referred to as "explainability" in deep learning parlance. One way to improve interpretability consists in adding layers based on morphological transforms [7]. Lately, progress have been made to develop a library called "morpho-layers" [2] that include morphological transforms and can be used in machine learning architectures [8]. Preliminary results have demonstrated the feasibility of this approach for the classification and regression problems on simulated datasets made of random point patterns, using "Choquet capacity" (max,+) layers [5].

Additionally, rigorous reconstruction theorems from random sets theory state how using such functional provides, in theory, complete statistical information [10].

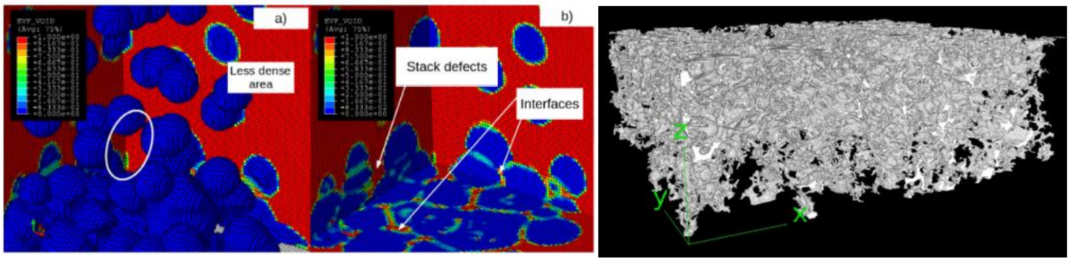
Another major difficulty is the limited number of experimental images, especially non-destructive, 3D dataset, such as micro-tomography images. They are often not sufficient to train neural networks with millions of parameters and designing neural network pipelines that include *a priori* information (geometrical percolation, particle size distribution, porosity) is a way to constraint generated microstructures into a meaningful space of approximation. From the perspective of stochastic geometry, many of these geometrical constraints can be described by the result of standard morphological transforms, such as the hit-or-miss transform (aka the Choquet capacity).

Objectifs / objectives of the PhD: The objectives of the PhD program are four-fold:

i. Design geometry/morphology-informed machine learning frameworks capable of generating complex binary microstructures and morphology, using morphological layers that can be interpreted such as the Choquet capacity;
ii. Assess the usefulness of the Choquet Capacity Networks with respect to standard machine learning frameworks in the context of a small number of images or limited datasets, and constrained by a priori knowledge, relevant to material processes in particular;
iii. Demonstrate the usefulness of such models to mimic experimental images of real materials ("interpolation problem");
iv. Generate microstructure images that are consistent with that of real materials, except for some features ("extrapolation problem").

Déroulement de la thèse / PhD program

The PhD thesis will consist in developing geometry/morphology-informed generative deep learning models. We will first address mathematically well-defined models, such as Boolean structures, based on Poisson point processes, with controlled features. A machine-learning architecture that makes use of the image transforms available in the morpho-layers library will be developed and leveraged to analyze synthetic data and structures, and compared to standard deep learning models. In contrast with classical machine-learning pipelines based on convolutional layers, the morpho-layer networks use mathematical morphology transformations and other layers implementing operators from stochastic geometry. When trained, the method amounts to learn "structural elements" that are at the basis of morphological operators (e.g. dilation, erosions). In doing so, the shape of these structural elements characterizes different grains morphology and spatial dispersion, for instance clusters and repulsion effects. From a theoretical point of view, these structural elements can be regarded as "input values" of the so-called Choquet-capacity functional [3]. Although, in practice, the functional cannot generally be computed beyond elementary characteristics (typically covariances) it has long been hypothesized that higher-order statistics could provide robust and accurate descriptors. The PhD thesis therefore aims to explore whether one may use machine-learning methods to provide relevant shape-descriptors, that are able to describe the spatial structure of random point processes with a good accuracy. The PhD thesis will focus as well on deep learning that use "small" parametric models and how to take into account geometrical constraints (relevant to processing methods). The ability of the models to generate complex microstructures from real materials will be assessed. Experimental images from cold-spray deposits as well as simulated data (see figure) will be studied. Finally, the PhD thesis will focus on how to generate microstructure images of the same type with different characteristics, such as grain granulometry curves that are different from that used to process the material.



Cold-spray Aluminium coating showing pore mechanisms formation at the particle scale (left; simulated impact, from [6]) and resulting large-scale pore distribution (right; tomography image of a real material).

References
[1] Molchanov, Ilya, and Ilya S. Molchanov. *Theory of random sets*. Vol. 19. No. 2. London: Springer, 2005.
[2] https://people.cmm.minesparis.psl.eu/users/velasco/morpholayers/intro.html
[3] Serra, J. (1981). The Boolean model and random sets. In Image modeling (pp. 343-370). Academic Press.
[4] Mohammadi, M., Velasco-Forero, S., Willot, F., Sangalli, M., & Angulo, J. (2023, June). Choquet Capacity Networks for Random Point Process Classification and Regression. In *International Conference on Continuum Models and Discrete Systems* (pp. 229-241). Springer Nature Switzerland.
[5] Weiller, S., Francesco, D., François, W., Alain, T., Jeandin, M., & Cédric, G. (2022). Influence of porosity on ultra-high vacuum gas-tightness in cold-sprayed aluminum coatings. *Transport in Porous Media*, *144*(2), 339-366.
[6] Weiller, S., & Delloro, F. (2022). A numerical study of pore formation mechanisms in aluminium cold spray coatings. *Additive Manufacturing*, *60*, 103193.
[7] Maragos, P., & Theodosis, E. (2019). Tropical geometry and piecewise-linear approximation of curves and surfaces on weighted lattices. *arXiv preprint arXiv:1912.03891*.
[8] S. Velasco-Forero, R. Pagès, J. Angulo, L. Learnable empirical mode decomposition based on mathematical morphology, SIAM Journal on Imaging Sciences 15 (1) (2022) 23–44.
[9] Chun, S., Roy, S., Nguyen, Y. T., Choi, J. B., Udaykumar, H. S., & Baek, S. S. (2020). Deep learning for synthetic microstructure generation in a materials-by-design framework for heterogeneous energetic materials. Scientific reports, 10(1), 13307.
[10] Kendall, D. G. (1974), Foundations of a theory of random sets, in E. F. Harding, D. G. Kendall, (eds), Stochastic geometry : a tribute to the memory of Rollo Davidson, Wiley, London, pp. 322–376.

Pertinence du projet au regard des missions et thématiques de l'ITN **:** Le projet de recherche permettra de (i) renforcer les travaux entrepris au CMM et au CMA dans le domaine de l'apprentissage profond, en particulier les liens entre morphologie, géométrie et apprentissage et la librairie Open Source *morpho-layers* développée au CMM ; (ii) de mettre en place un outil générique de génération sous contrainte de microstructures (dépôts ou autres) qui seront utilisés dans le cadre des travaux en cours par l'équipe "cold spray" (MAT). Nos travaux et résultats seront par ailleurs mis à disposition à la fois des chercheurs mais également des étudiants intéressés par les méthodes d'apprentissage appliquées à d'autres probl.

Profil de la candidate ou du candidat recherché(e) : nous recherchons un(e) excellent(e) étudiant(e), fortement motivé(e), en master 2 recherche ou équivalent à l'étranger, dans la thématique des sciences des données, analyse d'image ou de l'intelligence artificielle, possédant une expérience significative dans le domaine de l'apprentissage profond, un fort intérêt pour la recherche et les mathématiques en général.

Annexe : curriculum vitae des membres de l'équipe d'encadrement

Jesus Angulo https://people.cmm.minesparis.psl.eu/users/angulo/
François Willot https://people.cmm.minesparis.psl.eu/users/willot/tmp/willotCV.pdf
Santiago Velasco-Forero https://people.cmm.minesparis.psl.eu/users/velasco/
Francesco Delloro https://fr.linkedin.com/in/francesco-delloro-04549715b