Generation of cement paste microstructure using machine learning models

Liang, Minfei; Feng, Kun; He, Shan; Gan, Yidong; Zhang, Yu; Schlangen, Erik; Šavija, Branko

DOI
10.1016/j.dibe.2025.100624

Publication date
2025

Document Version
Final published version

Published in
Developments in the Built Environment

Developments in the Built Environment 21 (2025) 100624

Contents lists available at ScienceDirect

Developments in the Built Environment

journal homepage: www.sciencedirect.com/journal/developments-in-the-built-environment

Generation of cement paste microstructure using machine learning models

Minfei Liang [a,b], Kun Feng [c,*], Shan He [a], Yidong Gan [d], Yu Zhang [e], Erik Schlangen [a], Branko Šavija [a,**]

[a] Microlab, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, 2628CN, the Netherlands
[b] Clarendon Laboratory, Department of Physics, University of Oxford, Oxford, OX1 3PU, UK
[c] Key Laboratory of Transportation Tunnel Engineering, Ministry of Education, Southwest Jiaotong University, Chengdu, 610031, Sichuan, China
[d] School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China
[e] School of Civil Engineering, Southeast University, Nanjing, 211189, China

ARTICLE INFO

Keywords:
Cement paste
Microstructure
Generative deep learning
Denoising diffusion probabilistic model
Micromechanical analysis

ABSTRACT

The microstructure of cement paste determines the overall performance of concrete and therefore obtaining the microstructure is an essential step in concrete studies. Traditional methods to obtain the microstructure, such as scanning electron microscopy (SEM) and X-ray computed tomography (XCT), are time-consuming and expensive. Herein we propose using Denoising Diffusion Probabilistic Models (DDPM) to synthesize realistic microstructures of cement paste. A DDPM with a U-Net architecture is employed to generate high-fidelity microstructure images that closely resemble those derived from SEM. The synthesized images are subjected to comprehensive image analysis, phase segmentation, and micromechanical analysis to validate their accuracy. Findings demonstrate that DDPM-generated microstructures not only visually match the original microstructures but also exhibit similar greyscale statistics, phase assemblage, phase connectivity, and micromechanical properties. This approach offers a cost-effective and efficient alternative for generating microstructure data, facilitating advanced multiscale computational studies of cement paste properties.

1. Introduction

Concrete is the most widely used construction material in the world. As the main binder, cement paste plays a significant role in formulating the overall performance of concrete material. The properties of cement paste are governed by its microstructure. Compared to macroscale physical tests, studies on the microstructure provide a fundamental way to understand and quantify the material properties and behaviors (Bentz, 1999). A realistic microstructure is important input (Zhang et al., 2020) for many microscale modelling frameworks (e.g., continuum Finite Element Method (FEM) (Šmilauer and Bittnar, 2006; Rhardane et al., 2020), micromechanical homogenization schemes (Constantinides and Ulm, 2004; Liang et al., 2017; Gao et al., 2017), and discrete lattice models (Schlangen and Garboczi, 1997a, 1997b)), which aim to simulate physical properties of cement paste. The microstructural input can be obtained by physical experiments such as SEM microscopy and XCT, or hydration models such as Hymostruc (van Breugel, 1995; Van Breugel, 1993; Ye et al., 2003), μic (Bishnoi and Scrivener, 2009), and CemHyd3D (Bentz, 1997). On the one hand, the physical experiments can be time-consuming and expensive. On the other hand, due to the highly heterogenous nature of the microstructure and complexity of hydration reaction, the hydration models are subjected to certain simplifications and assumptions, such as the spherical shapes of cement particles and hydration products, and empirical coefficients in hydration kinetics calculations, etc.

The limitation mentioned above urges the need to efficiently obtain realistic microstructure input for the microscale computational frameworks (i.e., FEM, lattice model, etc.) of cement paste. The microstructure of heterogenous composites can be conceptualized as a complex probability distribution governed by random variables of very high dimensions. To efficiently reconstruct the microstructure, low-order probability functions have been used to approximate the probability function of the original microstructure using stochastic optimization techniques (Torquato, 2002a). Such methods have been utilized to reconstruct the concrete microstructure and calculate the properties such as permeability, stiffness, and thermal conductivity (Chung et al.,

* Corresponding author.
** Corresponding author.
E-mail addresses: minfei.liang@physics.ox.ac.uk (M. Liang), windfeng813@163.com (K. Feng), S.He-2@tudelft.nl (S. He), ygan@hust.edu.cn (Y. Gan), 101013462@seu.edu.cn (Y. Zhang), Erik.Schlangen@tudelft.nl (E. Schlangen), B.Savija@tudelft.nl (B. Šavija).

https://doi.org/10.1016/j.dibe.2025.100624
Received 21 December 2024; Received in revised form 6 February 2025; Accepted 11 February 2025
Available online 13 February 2025

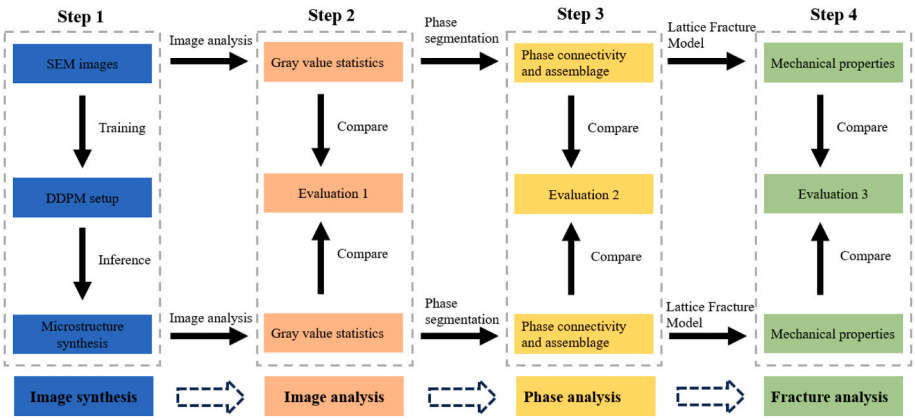

Fig. 1. Four-step workflow of this paper.

2014; Kim et al., 2019). However, due to the high computational cost, only two phases were considered in the low-order probability functions, which resulted in an incomplete description of the probability function of the original microstructure. Thus, discrepancy in morphology and macroscopic properties may still arise (Sahimi and Tahmasebi, 2021), even though the low-order probability functions of the synthesized and original data are identical.

Deep learning (DL) techniques have shown promise in extracting the latent representation from the microstructure of cement paste, which can then be used to predict the micromechanical properties (Liang et al., 2022, 2023a). More interestingly, generative DL such as variational autoencoders (VAE) (Kingma and Welling, 2013) and Generative Adversarial Networks (GAN) (Goodfellow et al., 2014) has been widely applied to synthesize complex microstructures in the last decade, described as below.

- The VAE uses an encoder-decoder structure to approximate the probability function of the microstructure with a mixed Gaussian distribution by maximum likelihood estimation. The encoder takes the microstructure as the input and represents the microstructure with a variational Gaussian distribution parameterized by its mean and variance. Then, the decoder samples from the variational distribution and generates the microstructure. The VAE has been used to generate two-phase steel and truss lattice metamaterials (Zheng et al., 2023a; Kim et al., 2021). However, when it comes to more complex datasets, such as high-resolution images, the assumption of Gaussian prior and posterior in VAE may result in low-quality and blurry images (Prince, 2023).
- Generative Adversarial Network (GAN) employs a generator-discriminator structure to generate microstructure through adversarial training. The generator takes random noise as input and generates microstructures via a deep neural network. The discriminator receives both the generated microstructures and real data as input and uses another deep neural network to classify them, determining whether they are real or generated. The generator and discriminator optimize alternately, continuously improving the quality of the generated microstructures during training (Goodfellow et al., 2014). GANs have been successfully applied to generate various microstructures, such as carbon steel and heterogeneous energetic materials (Safiuddin et al., 2021; Chun et al., 2020; Murgas et al., 2024).
- Recently, GANs have been successfully applied to synthesize cement paste microstructures. Assuming a four-phase microstructure of the cement paste, Hong et al. (2024) introduced an innovative CycleGAN-based framework for generating four-phase cement paste microstructures through a two-step process. In the first step, 2D slices were generated, which were subsequently used as input for another GAN to construct 3D microstructures. Their comprehensive statistical and micromechanical analyses demonstrated that the generated four-phase cement microstructures closely resembled the original ones. Similarly, Zhao et al. (2025) employed SliceGAN to generate 3D microstructures from 2D inputs, emphasizing systematic validation through in-depth microstructural and image analyses. This model bypasses the need for phase assumptions; however, it was noted that the fluctuating greyscale values in the microstructures may affect the consistency of the results. While these studies have significantly advanced the field, it is important to note that the reliance on 2D input data can introduce limitations in reflecting the true spatial complexity of 3D microstructures. Despite the successful application of GANs in generating microstructures of cement paste, there are some inherent limitations of the model. Training GANs can be notoriously unstable and requires careful tuning of hyperparameters and network architecture (Arjovsky and Bottou, 2017). Furthermore, GANs are prone to issues such as mode collapse, where the generator produces limited variations of outputs, and the vanishing gradient problem, which can hinder the training process (Metz et al., 2016).

Most recently, Denoising Diffusion Probabilistic Models (DDPM) (Ho et al., 2020) have emerged as a powerful alternative for generating microstructures, addressing some of the limitations inherent in VAE and GAN frameworks. DDPMs employ a forward process that gradually adds noise to the microstructure data, transforming it into a Gaussian distribution over several time steps. The reverse process then learns to denoise this Gaussian noise step-by-step, ultimately reconstructing the microstructure. Essentially, the iterative inference process of the DDPM can be viewed as the Hierarchical VAE (Vahdat and Kautz, 2020), an improved VAE which produces high-quality images by using iterative hierarchical Gaussian priors rather than single Gaussian priors. Therefore, DDPMs are less prone to generating blurry images compared to VAEs. In addition, the training of DDPMs tends to be more stable compared to GANs, avoiding issues like mode collapse and the vanishing gradient problem (Prince, 2023). DDPM is the basis of current most advanced image generators such as OpenAI's DALL-E2 and Google's Imagen. Based on DDPM, Lyu et al. (Lyu and Ren, 2024) reconstructed microstructures of various composite materials such as inclusion materials, spinodal decomposition materials, chessboard materials, fractal noise materials, etc. Similarly, Azqadan et al. (2023) generated microstructure images of distinct cast-forged AZ80 magnesium alloy components and obtained an average error of around 6.36% of the measured microstructural properties between the generated and original images.

In view of the advantages of DDPM over VAE and GAN, this paper employs the DDPM to synthesize the microstructural images of cement paste. The four-step workflow of this paper is shown in Fig. 1. In step 1, based on a database of 15000 SEM images, a DDPM will be constructed using a U-Net architecture, and then synthesized microstructures will be generated using the trained DDPM. Afterwards, image analysis (step 2),

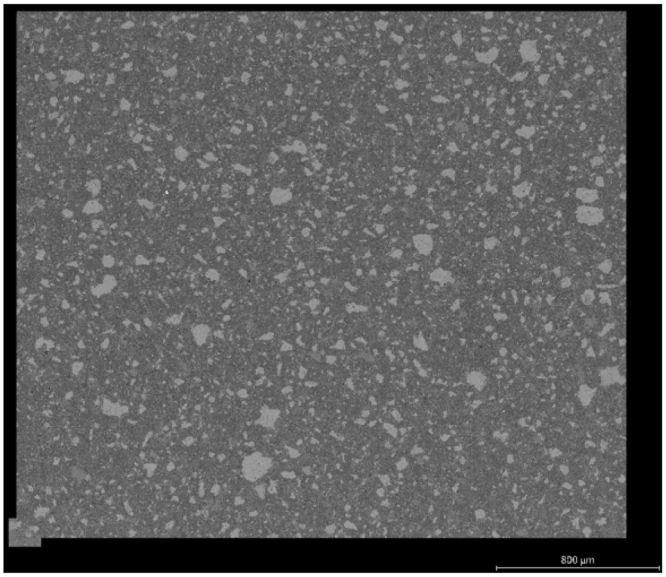

Fig. 2. BSE images for DDPM training
(Note: part of the image was used in our paper (Liang et al., 2023a) for predicting micromechanical properties).

phase segmentation (step 3) and fracture analysis (step 4) will be conducted to compare the synthesized and original microstructure.

2. DDPM for cement microstructure synthesis

This section first describes the database of SEM images for characterizing the microstructure of cement paste. Then, the theories, architecture, and training configurations of the DDPM used in this paper are presented.

2.1. Database

This paper uses backscattered electron (BSE) images to characterize the microstructure of cement paste and form a database for DDPM training. Herein, the sample preparation and image acquisition procedures will be introduced.

First, CEM I 52.5 Portland cement was used to prepare cement paste sample with a w/c ratio of 0.40 in sealed conditions. The cement paste was sealed and cured at a constant room temperature of 20 °C for 28 days. The solvent exchange method (Zhang and Scherer, 2011) was then used to stop the hydration with isopropanol solution. After arresting the hydration, the sample was cut into thin plate-like specimens using a micro-dicing saw to ensure parallel top and bottom surfaces. The sample preparation involved initial grinding with 4000 grit abrasive paper for 5 min, followed by polishing using a synthetic silk polishing cloth (MD-Dac from Struers) with 3 μm and 1 μm diamond paste in two separate 30-min sessions. An oil-based lubricant (DP-Lubricant Brown from Struers) was used during polishing to dissipate heat. To remove surface debris, the sample was immersed in an ultrasonic bath for 30 s between each grinding and polishing interval. The detailed information of the specimen preparation can be found in our previous paper (Liang et al., 2023a).

BSE images were taken using a FEI QUANTA FEG 650 SEM at an accelerating voltage of 15 kV. Prior to imaging, the sample was coated with a carbon layer approximately 10 nm thick. A 40 × 40 tile set covering the entire area of interest was prescribed, and the BSE images were taken automatically using commercial software (Maps 3 from Thermo Fisher Scientific) at a nominal magnification of 5000 × . These images were then electronically stitched into a single composite image (2.7 × 3.4 mm^2) as shown in Fig. 2, with a final resolution of 40 nm per pixel. Then, 15000 BSE images (100 × 100 μm^2) were randomly cropped from the composite image and resized to 128 × 128 pixels to form the database for DDPM training.

2.2. Denoising diffusion probabilistic model

DDPM comprises a forward diffusion and a reverse denoising process, as shown in Fig. 3. Both processes are Markovian, i.e., the state of each process at a certain step is only dependent on the state of the previous step. The forward diffusion process gradually adds noise to contaminate the image x_0 from step $t = 1 \sim T$, based on predefined transition function $q(x_t|x_{t-1})$, until the image finally becomes standard Gaussian noise. Then, the reverse denoising process aims to remove the noise step-by-step, using a neural network to approximate the denoising transition function $p_\theta(x_{t-1}|x_t)$. Note that both $q(x_t|x_{t-1})$ and $p_\theta(x_{t-1}|x_t)$ are assumed as Gaussian distribution. Essentially, DDPM can be viewed as hierarchical VAE since it assumes the noise in each step as Gaussian and generates the image by interactively predicting and removing the noise, while the VAE uses one-step mixed Gaussian distribution to approximate the original distribution and generate image. Therefore, the fitting capacity of DDPM is much less harmed by the assumption of "Gaussian distribution" in its transition functions, comparing to the one-step VAE.

In the forward diffusion process, the noise-adding procedure can be expressed as:

$$q(x_t|x_{t-1}) = N\left(x_t;\ \sqrt{1-\beta_t}\,x_{t-1}, \beta_t \boldsymbol{I}\right) \tag{1-1}$$

$$q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}) \tag{1–2}$$

where β_t is the predefined noise added to the image at step t. For $t = 1 \sim 1000$, β_t linearly increases from 0.0001 to 0.02. Leveraging the additivity of Gaussian distributions, the noised image at step t can be directly obtained from x_0 and β_t, as below:

$$q(x_t|x_0) = N(x_t;\ \sqrt{\bar{\alpha}_t}\,x_0, (1-\bar{\alpha}_t)\boldsymbol{I}) \tag{2}$$

where $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \prod_{i=0}^{t} \alpha_i$. Based on Bayes' rule and Eq (1) and (2), the posterior in the forward diffusion process can also be obtained, expressed as:

$$q(x_{t-1}|x_t) = N(x_{t-1};\ \tilde{\mu}_t(x_t, x_0), \tilde{\beta}_t \boldsymbol{I}) \tag{3-1}$$

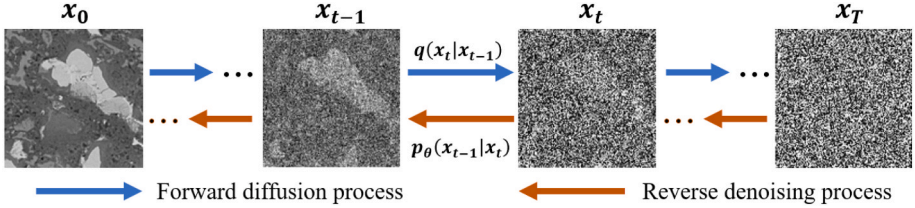

Fig. 3. The forward diffusion and reverse denoising process of DDPM.

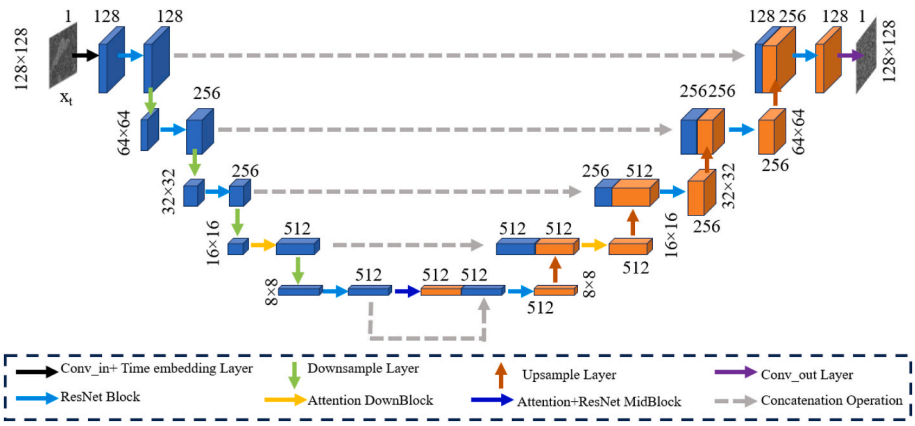

Fig. 4. U-Net architecture used for noise prediction in reverse denoising process.

$$\widetilde{\mu}_t(x_t, x_0) = \frac{\sqrt{\overline{\alpha}_{t-1}}\beta_t}{1-\overline{\alpha}_t}x_0 + \frac{\sqrt{\alpha_t}(1-\overline{\alpha}_{t-1})}{1-\overline{\alpha}_t}x_t \tag{3-2}$$

$$\widetilde{\beta}_t = \frac{1-\overline{\alpha}_{t-1}}{1-\overline{\alpha}_t}\beta_t \tag{3-3}$$

The transition function in the reverse denoising process is also a Gaussian distribution, as below:

$$p_\theta(x_{t-1}|x_t) = N\left(x_{t-1};\ \mu_\theta(x_t, t), {\sigma_t}^2 \boldsymbol{I}\right) \tag{4}$$

where ${\sigma_t}^2$ can be assumed the same as β_t or $\widetilde{\beta}_t$, which both lead to similar results (Ho et al., 2020). In this paper, we set ${\sigma_t}^2 = \widetilde{\beta}_t$. Then, the loss function of the training process is defined based on maximum likelihood estimation. Similar to VAE, the training is formulated to optimize the evidence lower bound (ELBO) on the negative log likelihood, as below:

$$\mathbb{E}[-\log p_\theta(x_0)] \le \mathbb{E}_q\left[-\log\frac{p_\theta(x_{0:T})}{q(x_{1:T}|x_0)}\right] = \mathbb{E}_q\Bigg[- \log p(x_T) - \sum_{t\ge 1}\log\frac{p_\theta(x_{t-1}|x_t)}{q(x_t|x_{t-1})}\Bigg] \tag{5}$$

Rewriting Eq (5) and then the loss function can be obtained, as below:

$$L = \mathbb{E}_q[D_{KL}(q(x_T|x_0)\|p(x_T)) - \log p_\theta(x_0|x_1) + D_{KL}(q(x_{t-1}|x_0, x_t)\|p_\theta(x_{t-1}|x_t))] \tag{6}$$

where D_{KL} is the Kullback–Leibler divergence (Goodfellow et al., 2016), which is a common measure of distance between two probability distributions and can be calculated by $D_{KL}(p(x)\|q(x)) = \int p(x)\log\frac{p(x)}{q(x)}dx$. The first term in Eq (5) is a constant term and can be ignored, since both $q(x_T|x_0)$ and $p(x_T)$ are standard Gaussian distribution. The second term is the reconstruction term. The last term is the distance between the posterior of the forward diffusion (Eq (3)) and the transition in the reverse process (Eq (4)). Because we set ${\sigma_t}^2 = \widetilde{\beta}_t$, minimizing the last term is equivalent to fitting $\mu_\theta(x_t, t)$ with $\widetilde{\mu}_t(x_t, x_0)$. By reparametrizing the second and third term following Eq (3), a simplified loss function for the DDPM can be obtained, as below:

$$L_{sym} = \mathbb{E}_{t,x_0,\epsilon}\left[\left\|\epsilon - \epsilon_\theta\left(\sqrt{\overline{\alpha}_t}x_0 + \sqrt{1-\overline{\alpha}_t}\epsilon, t\right)\right\|^2\right] \tag{7}$$

where $\epsilon_\theta\left(\sqrt{\overline{\alpha}_t}x_0 + \sqrt{1-\overline{\alpha}_t}\epsilon\right)$ is the neural network which takes the original images x_0 and step t as the input and predicts the stepwise noise that needs to be removed; ϵ is the noise sampled from a standard Gaussian distribution in the forward diffusion process.

2.3. U-net architecture for noise prediction

This paper uses an U-Net architecture (Ronneberger et al., 2015) for the noise prediction in Eq (7), as shown in Fig. 4. A classic encoder-decoder structure with symmetrical downsampling and upsampling paths is used to capture and reconstruct the multiscale image features. The input block is composed of a convolutional layer and time-embedding layer. This time-embedding layer uses sinusoidal positional encodings (Vaswani et al., 2017) to convert the timestep into a high-dimensional vector, which is then processed through linear layers to integrate time step information into the model.

In both the downsampling and the upsampling paths, five downsampling and upsampling blocks are used respectively. Each block consists of a ResNet block (He et al., 2015) followed by a downsample or an upsample layer. ResNet block contains group normalization layers (Wu and He, 2018), residual convolutional layers, time projection layers, and Silu activation functions (Elfwing et al., 2017). The time projection layers takes the time step vector directly from the input layer and project it to a new tensor matching the feature tensor size. Then, the time step tensor is added to the feature tensor to incorporate the time step information. The downsample layer uses a convolution with a stride of 2. The upsample layer uses a transposed convolution and nearest-neighbor interpolation. Skip connections are used between downsampling and upsampling blocks, where feature maps are concatenated to preserve spatial information. Attention (Vaswani et al., 2017) layers are added both in the downsampling and upsampling paths to capture the global context and enhance feature extraction. A mid block is positioned between the downsampling and upsampling paths, which comprises ResNet layers and attention layers.

2.4. Training and sampling

The U-Net is trained in mini-batches with a batch size of 32 and a total training iteration number of 48000. Prior to training, all input images, originally comprising grey values within the range [0, 256], are first normalized to the [0, 1] interval by dividing each value by 256. Subsequently, these normalized values are linearly scaled to the range $[-1.0, 1.0]$ through the transformation $x' = 2x - 1$. The maximum step T of the forward diffusion process is 1000. In the training process, the time step t is sampled from a uniform distribution of $\{1, 2, 3, \ldots, T\}$, and the noise ϵ is sampled from a standard Gaussian distribution. Using the original image x_0, time step t, and the noise ϵ as input, the diffusion process (Eq (2)) can be conducted to get the noised image x_t. The U-Net can then take the noised image x_t and time step t as the input to predict the noise, i.e., ϵ_θ, for the reverse denoising process. Using the gradient descent optimizer Adam (Kingma and Ba, 2014), the parameters of the U-Net (i.e., θ) are optimized according to the loss function in Eq (7). In the gradient descent optimization, a cosine warm-up schedule for the

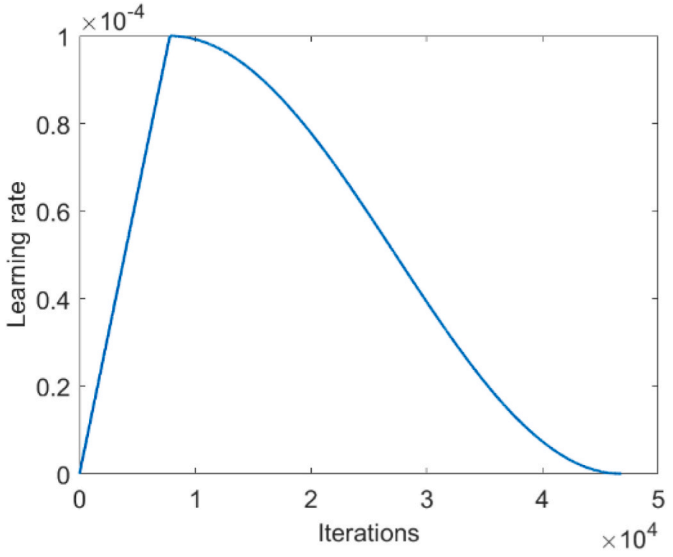

Fig. 5. The cosine warm-up schedule of the learning rate.

learning rate is used, which comprises a linear warm-up in the first 8000 iterations and a cosine-shaped decay afterwards, as shown in Fig. 5.

Afterwards, new microstructures can be sampled. First, sample x_T from a standard Gaussian distribution and then conduct the reverse denoising process to get the denoised image x_{T-1} using the trained U-Net. By repeating such a denoising process from step T to 0, a new microstructure can be obtained. In Fig. 6, the reverse denoising process for generating new microstructures is shown for four cases at five representative steps: $x_{1000}, x_{300}, x_{200}, x_{100}, x_0$. The x_{1000} is a standard Gaussian noise; $x_{300}, x_{200}, x_{100}$ are partially-denoised microstructure; and the x_0 is the final microstructure. Fig. 7 presents ten original and generated microstructures, respectively. The detailed comparison between the original and generated microstructures will be conducted by comprehensive image analysis, phase segmentation, and lattice fracture simulation afterwards.

3. Evaluation framework

An evaluation framework is built up herein to compare the generated microstructures with the original microstructures. It is important to note that the DDPM used in this study is not designed to replicate specific original images. During training, the U-Net learns to predict the noise in each reverse denoising step (Fig. 3). In the generation phase, 1000 denoising steps are applied to synthesize new microstructures starting from standard Gaussian noise. As a result, there is no one-to-one correspondence between generated and original images and pairwise comparisons are not feasible. Instead, we conducted comprehensive dataset-level comparisons to evaluate the global statistical and micromechanical characteristics of the generated microstructures. The dataset for evaluation analysis comprises 1280 generated microstructures and 1280 original microstructures. Based on the evaluation dataset, the evaluation analysis is conducted in the following three steps.

1) Extract and compare the statistical characteristics of the greyscale values of generated and original images in the evaluation dataset, i.e., the probability density distribution (PDF) and the cumulative density distribution (CDF).
2) Based on the PDF and CDF, conduct phase segmentation on the evaluation dataset, calculate and compare the phase assemblage of the original and generated microstructures, and quantitatively analyze the microstructure characteristics.
3) Based on the segmented microstructures, conduct micromechanical simulations on the whole evaluation dataset, and compare the micromechanical properties and behaviors (i.e., elastic modulus, uniaxial tensile strength, and stress-strain curves).

The procedures above aim to provide a comprehensive description of the difference between the generated and original microstructures. The first step of PDF/CDF analysis is straightforward. The theories and methods for phase assemblage analysis and micromechanical analysis are introduced in this section and the results are presented in section 4.

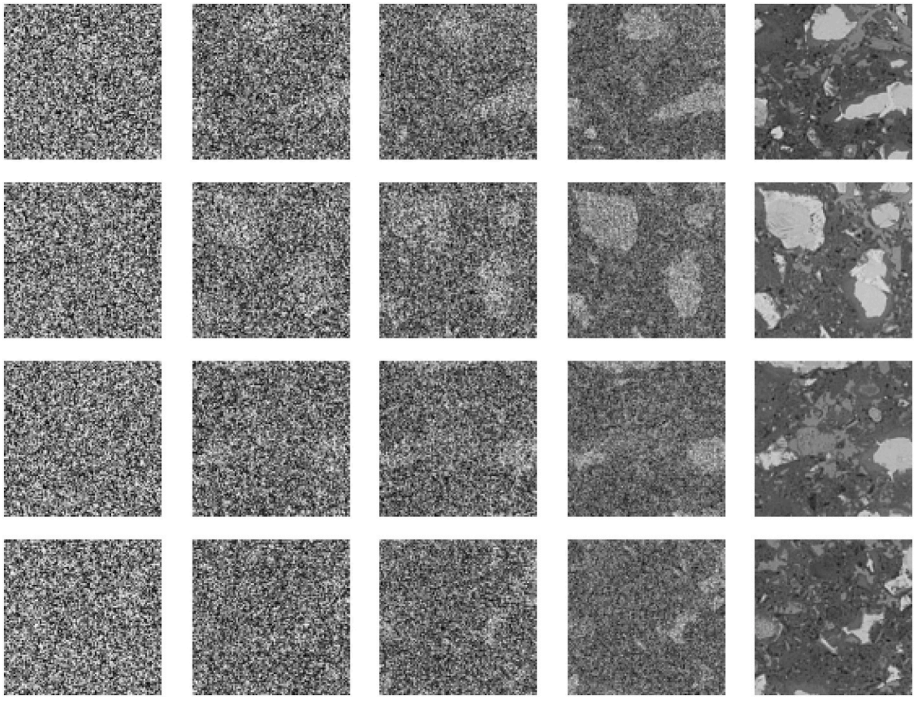

Fig. 6. Four cases of the reverse denoising process to generate new microstructures
(Note: Images from the top to bottom row are four different cases; Images from the left to right column corresponds to: $\boldsymbol{x_{1000}}, \boldsymbol{x_{300}}, \boldsymbol{x_{200}}, \boldsymbol{x_{100}}, \boldsymbol{x_0}$)

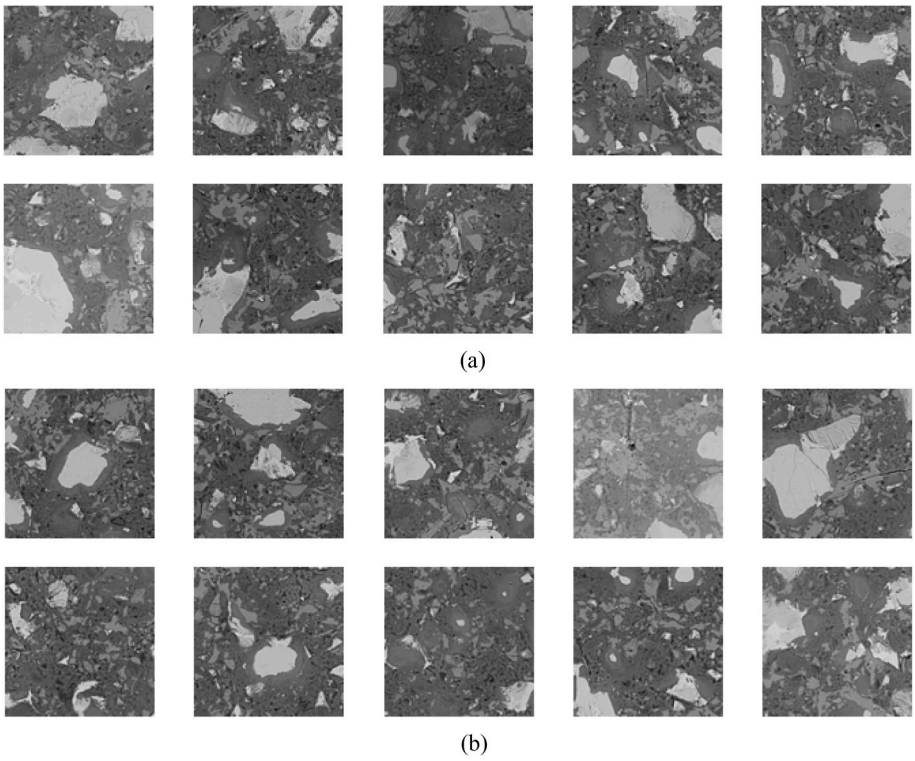

Fig. 7. Ten cases for comparing the original and generated microstructures: (a) Cases 1–5; (b) Cases 6–10. (Note: The top and bottom rows correspond to the original and generated microstructures respectively.)

3.1. Phase assemblage analysis

The brightness in greyscale of the BSE images is primarily a function of the atomic number in the area of interest (Scrivener et al.), which can be used to represent the microstructural characteristics. Based on the PDF and CDF of the greyscale values, the microstructure of cement paste can be typically divided in four different phases, i.e., pores, low-density calcium-silicate-hydrate (LD- CSH), high-density calcium-silicate-hydrate (HD-CSH), and unhydrated cement (UHC) (Zhang et al., 2018a; Constantinides and Ulm; Bernard et al., 2003). The threshold value for phase segmentation can be calculated by global thresholding method (Wong et al., 2006). The mass ratio between the LD-CSH and HD-CSH M_r can be calculated by Jennings' empirical model (Tennis and Jennings, 2000), as below:

$$M_r = 3.017\left(\frac{w}{c}\right)\alpha - 1.347\alpha + 0.538 \tag{8}$$

where $\frac{w}{c}$ is the water/cement ratio; α is the hydration degree and can be calculated as below:

$$\alpha = \frac{\frac{V_h}{V}}{\frac{V_h}{V} + V_u} \tag{9}$$

where V_h is the volume fraction of hydration product (i.e., LD-CSH and HD-CSH); V_u is the volume fraction of UHC; V is the volume ratio of hydration product to reactant, which is assumed as 2.2 according to (Van Breugel, 1993). Based on global thresholding method (Wong et al., 2006) and Eq (8) and (9), the threshold values of the four phases including pores, LD-CSH, HD-CSH, and UHC can be calculated and therefore the BSE images can be segmented.

3.2. Phase connectivity analysis

This paper characterizes the connectivity of microstructural phases by lineal path function (LPF) (Torquato, 2002b), which was proposed to quantify the connectivity of microstructural phases. The computation of LPF herein is conducted using the open-source PoreSpy library (Gostick et al., 2019). By inserting lines of arbitrary length into the microstructures, the LPF quantifies the connectivity based on the probability of lines with specific length that lie wholly within a single phase. LPF has been successfully used to compare the statistical identity of cementitious microstructures (Hong et al., 2024; Han et al., 2018).

As will be illustrated in the micromechanical analysis and by other relevant studies (Hong et al., 2024; Zhang et al., 2019a; Han et al., 2023), cracks are more prone to propagate through the weak phases of the LD-CSH (or outer product) and pore and therefore these two phases primarily dominate the micromechanical performance. In this regard, this paper computes the LPF of pore and LD-CSH for all the generated and original microstructures within the evaluation dataset, which can be

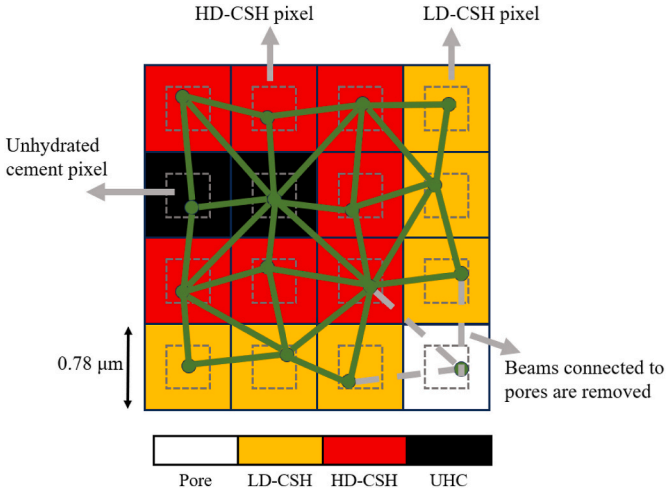

Fig. 8. Lattice network for micromechanical analysis of cement paste.

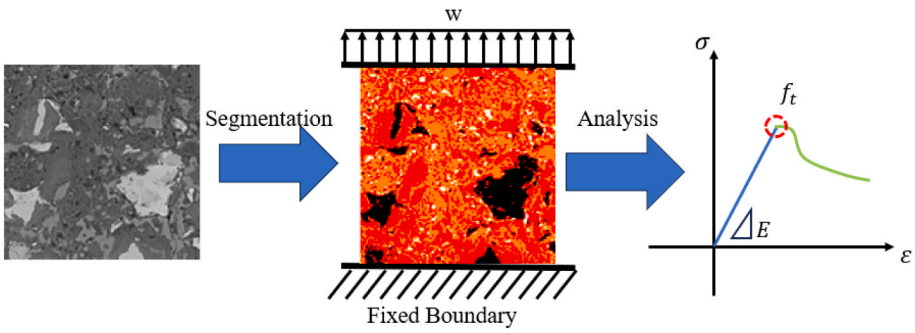

Fig. 9. Boundary conditions and result analysis of micromechanical analysis.

Table 1
Local properties of different phases.

Phases	f_i/MPa	E/GPa
UHC	614.7	84.2
LD-CSH	52.2	21.3
HD-CSH	82.8	26.4

essential statistics to evaluate the effectiveness of the proposed DDPM model.

3.3. Micromechanical analysis

The micromechanical analysis is conducted by the Delft Lattice Model (Schlangen and Garboczi, 1997a), which has been widely applied in simulating various micromechanical behaviors and validated by series of micromechanical tests (Zhang et al., 2019b; Gan et al., 2020; Chang et al., 2022; Šavija et al., 2013a, 2013b). For the micromechanical analysis, consistent with the phase assemblage analysis described above, we assumed that the cement paste is a four-phase composite that is comprised of pore, LD-CSH, HD-CSH, and UHC. Such four-phase assumption was used for lattice model of micromechanical analysis in our previous paper (Liang et al., 2022; Gan et al., 2021). Note that different assumptions of phases may be used for the model. In this paper, we use the four-phase assumption for the micromechanical analysis of both original and generated images and compare their differences.

Using the segmented microstructure of cement paste, a Lattice network can be established with different local micromechanical properties for different phases, as shown in Fig. 8. For each microstructural input, a 128 × 128 array of square cells is generated, and each cell is assigned with different element types. Then, sub cells with the side length equal to half of the square cells (Zhang et al., 2018b) are generated in the center. Within each sub cell, the nodes of the lattice beam elements are randomly positioned following a uniform distribution. Finally, based on the random beam nodes array, Delaunay triangulation is performed to connect the neighboring nodes and form the lattice network (Šavija et al., 2014).

Afterwards, based on the segmentation results, virtual uniaxial tensile tests are conducted on the microcubes, by fixing the bottom side and applying tensile stress on the top side, as shown in Fig. 9. The specimen is loaded until breaking, then the elastic modulus E (i.e., the slope of the loading phase) and tensile strength f_t (i.e., the peak) can be obtained from the stress-strain curve. The local properties of different phases are determined according to micromechanical testing results. As shown in Table 1, the values of elastic modulus for different phases are obtained from deconvolution results of statistical nanoindentation tests (Hu and Li, 2014). The values of tensile strength for different phases cannot be tested and was obtained by a numerical calibration procedure comparing the modelling results of lattice model and the results of microscale cantilever bending tests (Gan et al., 2021). In this paper, the resolution of each square cell is 0.78 μm, which is less than 5 μm and therefore the local properties of each phase are not dependent on w/c ratios, according to (Hu and Li, 2015; Vandamme and Ulm, 2013).

The beam elements in the lattice network are assumed as Timoshenko beams with circular cross-sections and the element stiffness matrix k can be calculated accordingly. Afterwards, the global stiffness matrix K can be assembled. The fixed boundary condition is achieved by the penalty method, i.e., by assigning a large value (10^6 GPa) for stiffness on the bottom side. The lattice model is then solved based on the linear equations of Timoshenko beam system $S = F/K$, where S is the nodal displacement and F is the nodal force on the top side. The loading of the specimen is conducted in 200 load-unload steps. In each step, a beam element that meets the maximum stress criterion will be removed, which characterizes the damage induced by the immediate constant load. According to (Zhang et al., 2017; Schlangen and van Mier, 1992), the stress in each lattice element is calculated as follows:

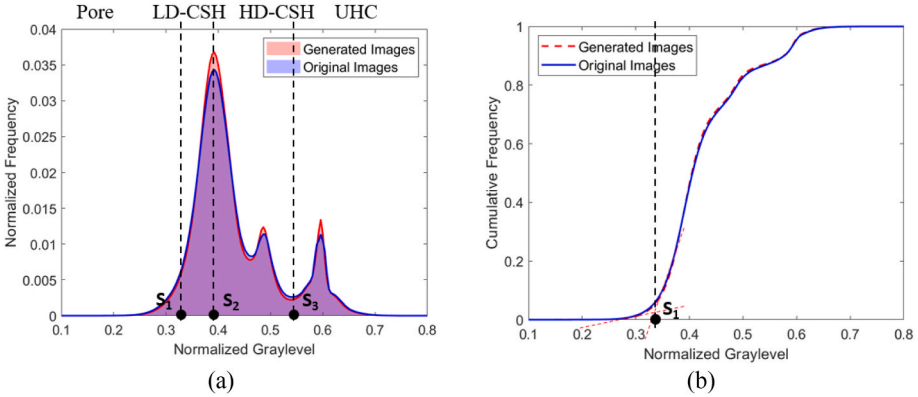

Fig. 10. Greyscale statistics of the generated and original images: (a) PDF; (b) CDF.

$$\sigma_f = \alpha_N \frac{N}{A} + \alpha_M \frac{M}{W} \tag{9}$$

where A is the cross-sectional area of every beam element; W is the cross-sectional moment of resistance of every beam element ($W = \pi D^3/32$, where D is the effective diameter of the lattice element.); N and M are the axial force and bending moment of the beam element; α_N and α_M are the influencing factors of axial force and bending moment, which normally equal 1.0 and 0.05, respectively, according to (Qian, 2012; Šavija et al., 2020; Chang et al., 2020).

4. Results and discussion

This section presents a comparison of the generated and original microstructures by evaluating the greyscale statistics, phase assemblage, and micromechanical properties. Suggestions for future work are discussed at the end.

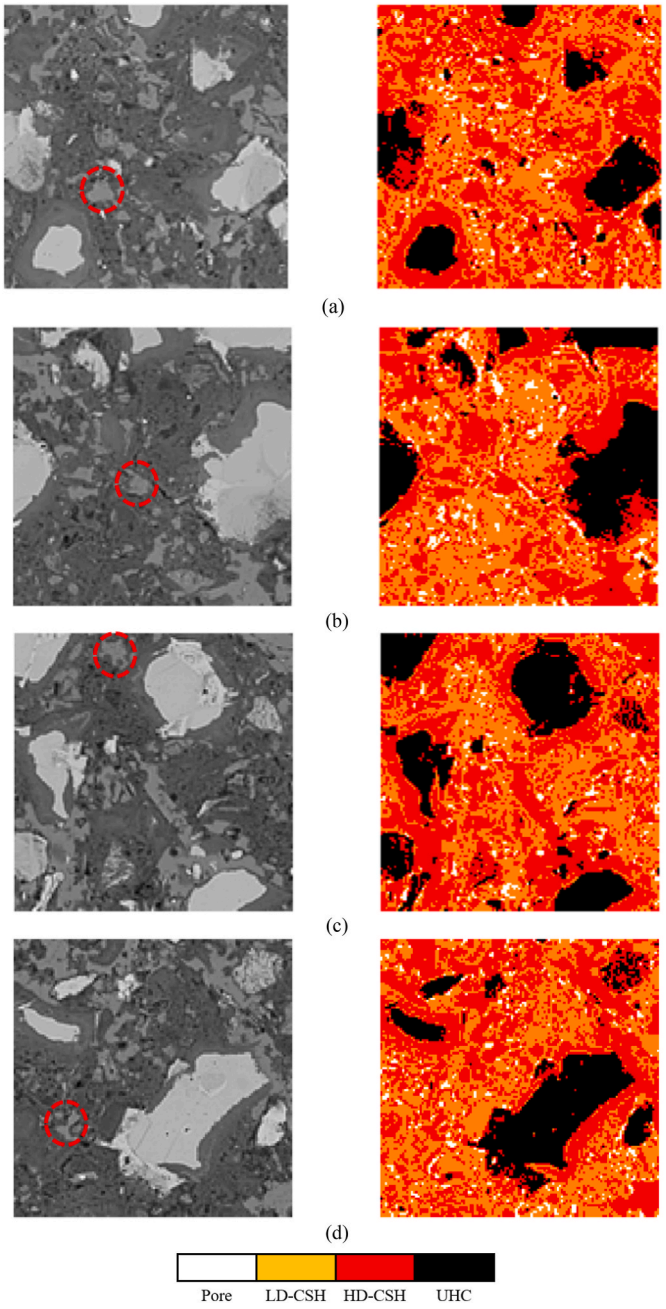

Fig. 11. Four representative cases of phase segmentation: (a–b) Generated microstructures (c–d) Original microstructures. (Note: Red dashed circles mark representative CH clusters).

4.1. Image analysis

Based on the evaluation dataset (which comprises 1280 generated images and 1280 original images), the PDF and CDF of the greyscale values of the generated and original images can be extracted respectively, as shown in Fig. 10. The PDF and CDF of the generated and original images highly resemble each other. The CDFs are almost identical. In the PDFs, the locations of three peaks are almost the same for both generated and original images. Slight differences can be found regarding the frequency of the PDF, especially at the peaks. Overall, the comparison shows that DDPM is able to generate BSE images of cement paste with the same overall distribution of greyscale values.

4.1.1. Segmentation

Furthermore, the three threshold values S_1, S_2, S_3 for segmenting the microstructures can be calculated based on the PDF and CDF by the global threshold method introduced in section 3.1. Greyscale value between $0 \sim S_1$ is range of pore phase and the upper threshold value S_1 can be determined from the inflection point of the CDF (Fig. 10(b)) (Wong et al., 2006). S_2 and S_3 are the upper threshold values of LD-CSH and HD-CSH, respectively. Note that the second peak of PDF corresponds to the calcium hydroxide (CH). Despite the distinct peak of CH, the segmentation of the CH phase remains a challenge for BSE images because of its complex interpenetrations with other phases (Scrivener, 2004). In this paper, we defined the HD-CSH a composite mainly comprised of inner product and CH, according to (Zhang et al., 2020; Vandamme, 2002). The four-phase assumption used in this paper was well identified in statistical nano-indentation tests (Liang et al., 2023b; Hu et al., 2016). The results of S_1, S_2, S_3 for the generated images are 0.332, 0.401, and 0.544, while the ones for the original images are 0.329, 0.405, and 0.546.

4.1.2. Phase assemblage

Segmenting the generated and original images using the threshold values above, the four-phase microstructures can be obtained. To qualitatively show the segmentation results, four representative cases of segmentation are shown in Fig. 11. A common understanding over the microstructure of cement paste can already be seen in both the generated and original images: The CSH can be categorized as HD-CSH and LD-CSH, or the so-called inner product and outer product (Taylor, 1997). The HD-CSH (i.e., inner product) is produced in the space originally occupied by UHC. Therefore, the HD-CSH should precipitate around the UHC and form a rim, as shown in the microstructure of the original images (Fig. 11(c–d)). Interestingly, such HD-CSH rim around the UHC also clearly forms in the generated microstructures. It should also be noted that, despite the four-phase assumption include CH in the HD-CSH, the distribution pattern of CH in both the generated and original images is still distinguishable. According to (Scrivener et al., 2016), the CH clusters (identified as HD-CSH) is produced in the water-filled space and mixed with outer product (identified as LD-CSH). Representative CH clusters are highlighted with red dashed circles in Fig. 11 to illustrate their distribution. Accordingly, in the generated images we can also see the intertwined distribution of LD-CSH and CH clusters distributing in the water-filled space. The distribution pattern of LD-CSH and HD-CSH in the generated images indicates that the DDPM successfully captures such microstructural features.

However, it should be noted that cementitious microstructures exhibit significant variability from sample to sample. For a quantitative and reliable evaluation, all microstructures in the evaluation dataset were segmented, and key indices—including porosity, UHC ratio, degree of hydration (DOH), and the volume ratio of LD-CSH—were calculated for both generated and original samples. The mean values and standard

Table 2
Overall microstructural indices of generated and original microstructures.

Indices	*Generated*	*Original*
Porosity (%)	4.10 ± 1.70	4.61 ± 1.80
LD-CSH ratio (%)	42.20 ± 8.18	40.92 ± 8.04
UHC ratio (%)	12.90 ± 8.56	13.22 ± 8.44
DOH	0.757 ± 0.125	0.750 ± 0.123

deviations of these indices are presented in Table 2. The close statistical alignment between the generated and original datasets indicates that the model not only replicates the average microstructural properties but also effectively captures the inherent variability. The relatively large standard deviations reflect the natural microstructural heterogeneity of cementitious materials. Nevertheless, the generated microstructures show slightly higher DOH, more LD-CSH, and lower porosity compared to the original (i.e., real) microstructures.

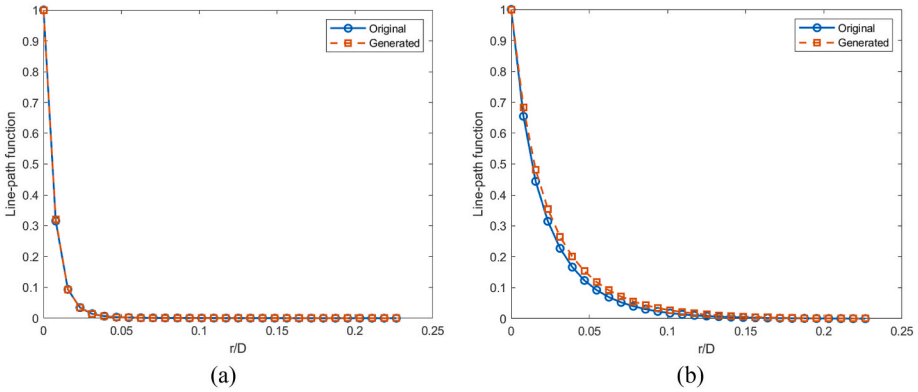

Fig. 12. LPF function calculated for pore (a) and LD-CSH (b), averaged over generated and original microstructures respectively within the evaluation dataset. (Note: r is the line length and D is the width of each microstructure).

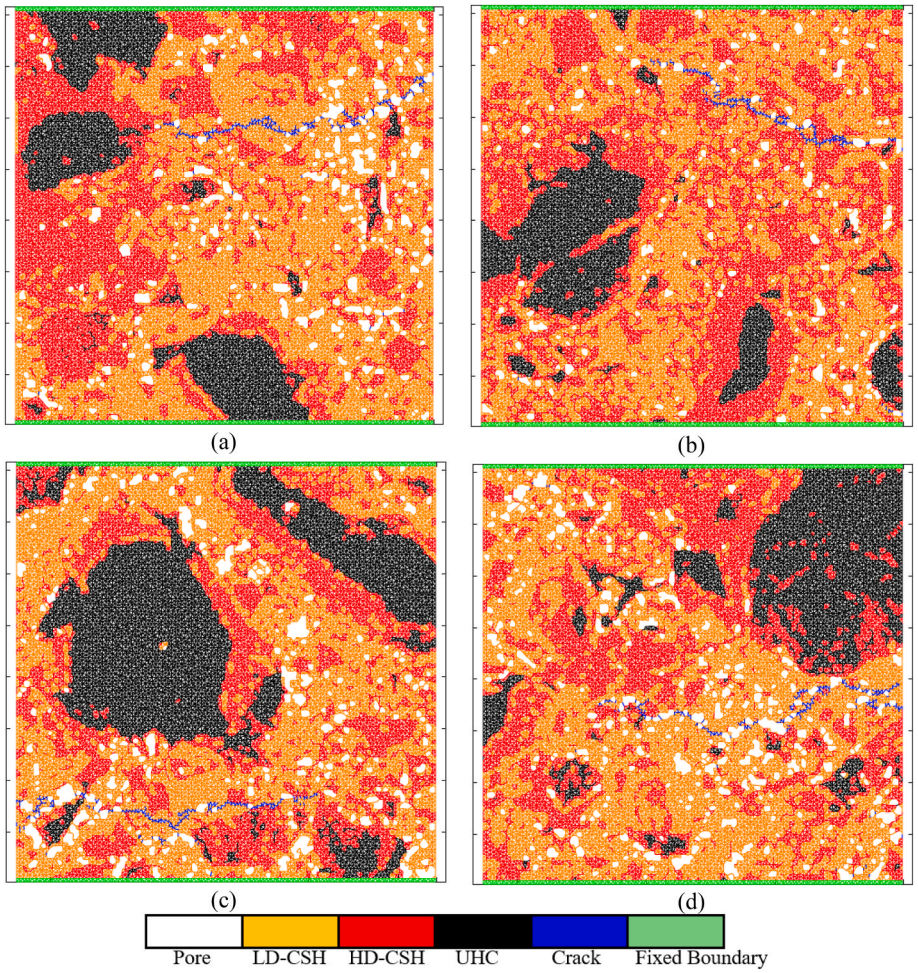

Fig. 13. Four representative cases of cracking patterns: (a–b) Generated microstructures (c–d) Original microstructures.

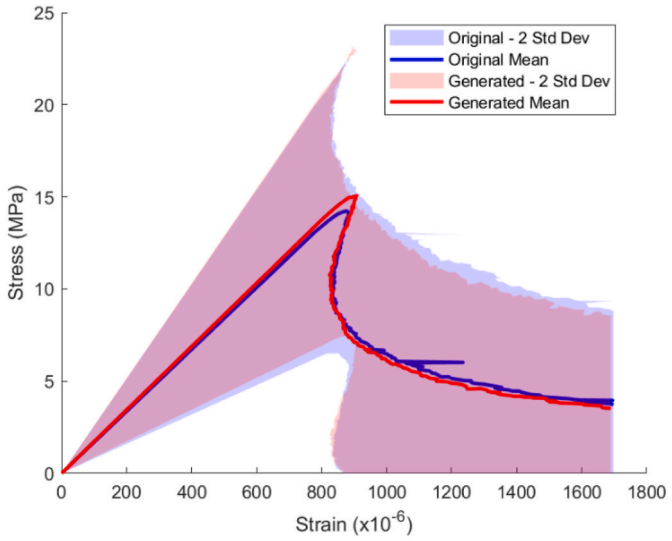

Fig. 14. Stress-strain curves.

4.1.3. Phase connectivity

Using the LPF, this paper quantifies the connectivity of the pore and LD-CSH phases. The LPF is calculated for all generated and original microstructures within the evaluation dataset along both the x- and y-axis. The CDF of the averaged LPF over the dataset is shown in Fig. 12. In general, the LPF results show strong similarity over pore and LD-CSH connectivity between generated and original microstructures. Naturally, the LPF distribution of LD-CSH covers wider range in line length (0∼0.1 r/D) while the LPF of pores only concentrates over the range of 0∼0.03 r/D, reflecting the inherent size differences between these phases. A slight discrepancy was observed in the LD-CSH phase, where the generated microstructures exhibited marginally better connectivity. This aligns with the phase assemblage analysis, where generated microstructures showed a slightly higher LD-CSH ratio (Table 2).

To sum up, this section shows that the generated images by the DDPM resemble the original microstructures characterized by BSE-SEM. The high similarity on the greyscale statistics (Fig. 10), the indices (Table 2), and phase connectivity (Fig. 12) shows that DDPM can well capture the global microstructural characteristics. Moreover, the generated images also indicate similar local microstructural characteristics, that the HD-CSH mostly forms as a rim surrounding the UHC particles. Nevertheless, such indication cannot give quantitative conclusions. To this end, micromechanical analysis needs to be conducted, whose results are determined by not only global but also local micromechanical characteristics.

4.2. Micromechanical analysis

Using the Delft Lattice Model for micromechanical analysis, this section aims to quantify the similarity between the generated and original microstructures by qualitatively comparing the cracking patterns and quantitatively comparing the stress-strain curves, elastic modulus, and uniaxial tensile strength. Four representative cases of the cracking patterns at the last simulation step are shown in Fig. 13. The results show that cracks mostly propagate through the LD-CSH, which conforms to the classical understandings of the LD-CSH formation from the aspect of cement chemistry (Taylor, 1997): LD-CSH (i.e., outer product) forms in the water-filled pore space. During hydration, because the volume of hydration products is smaller than that of the reactants, chemical shrinkage happens and creates pores. Therefore, compared with HD-CSH, LD-CSH is a more porous region and therefore the main cracks propagate through LD-CSH. Both generated and original microstructures are dominated by such cracking patterns, indicating that the DDPM captures the fact that LD-CSH has a more porous microstructure and therefore is most probable to allow for crack propagation in fracturing tests.

The stress-strain curves of the virtual uniaxial tensile tests of the whole evaluation dataset are shown in Fig. 14. The results show that they have almost the same average stress-strain curve, while the standard deviation regions almost coincide. Nevertheless, the generated microstructures tend to have slightly higher uniaxial tensile strength, which is expected if considering the generated images also show slightly higher DOH and lower porosity, see Table 2.

Extracting the uniaxial tensile strength and elastic modulus from the stress-strain curves, the statistical results of the whole evaluation dataset are shown in Fig. 15. The histogram (Fig. 15(a and b)) illustrates that the distributions of both strength and elastic modulus in the original and generated microstructures visually resemble Gaussian distributions. The quantile-quantile (QQ) plot Fig. 15(c and d) compares the theoretical quantiles of a standard Gaussian distribution (x-axis) with the actual sample values (y-axis) of elastic modulus and strength. Data points closely following the 45-degree reference line indicate a strong agreement with Gaussian behavior. The P values are also calculated using Kolmogorov–Smirnov (KS) test. All the p values are higher than 0.05, further confirming the observation that both the elastic modulus and the uniaxial tensile strength of generated and original microstructures follow a Gaussian distribution. It is worth noting that the P values are highly sensitive to slight variations in the distribution tails, which may account for differences observed in KS tests.

The PDFs Fig. 15(e and f) are estimated by fitting the sample data with a kernel smoothing function. Herein a Gaussian kernel is used for the fitting process. Both the results of elastic modulus and strength show almost the same distribution for the generated and original microstructures. Comparing the mean values of the generated and original microstructures, the strength and elastic modulus are 17.13 MPa vs. 16.26 MPa and 16.31 GPa vs. 15.88 GPa, with a difference of 5.4% and 2.7%, respectively, indicating good consistency between the simulated micromechanical properties of generated and original images. Similarly, it is found here that the generated microstructures have higher strength and elastic modulus, which is consistent with the findings in stress-strain curves and in DOH and porosity indices in phase assemblage analysis. The difference is somewhat larger in uniaxial strength. This is because that the uniaxial tensile strength is highly influenced by the weak areas, i.e., the porous LD-CSH, as explained in Fig. 13. The randomness of local pore sizes and distributions dominates the fracture process and therefore results in variance of the strength data (see Fig. 15(b–d)).

To summarize, this section presents evidence that the generated images have almost the same micromechanical behaviors/properties the original images, based on comprehensive micromechanical analysis over the evaluation dataset. Since micromechanical behaviors/properties are generally a combined outcome of multiple microstructural effects (Hu et al., 2017; Ulm et al., 2010), this section provides strong evidence that the generated images are identical to the original images.

4.3. Discussion

This paper presents the application of DDPM in synthesizing realistic microstructures of cement paste directly using the BSE images as training input. The microstructures generated by the DDPM do not only visually resemble the original ones, but also show almost identical greyscale statistics, phase assemblage and micromechanical properties, indicating that the DDPM is able to capture the underlying micromechanical representation. Without any assumptions and solely informed by the BSE images, the DDPM can generate realistic microstructure input for microscale studies of cement paste.

To extend the work of this paper to build a more general, versatile, and useful tool for the cement and concrete research, the following works are recommended.

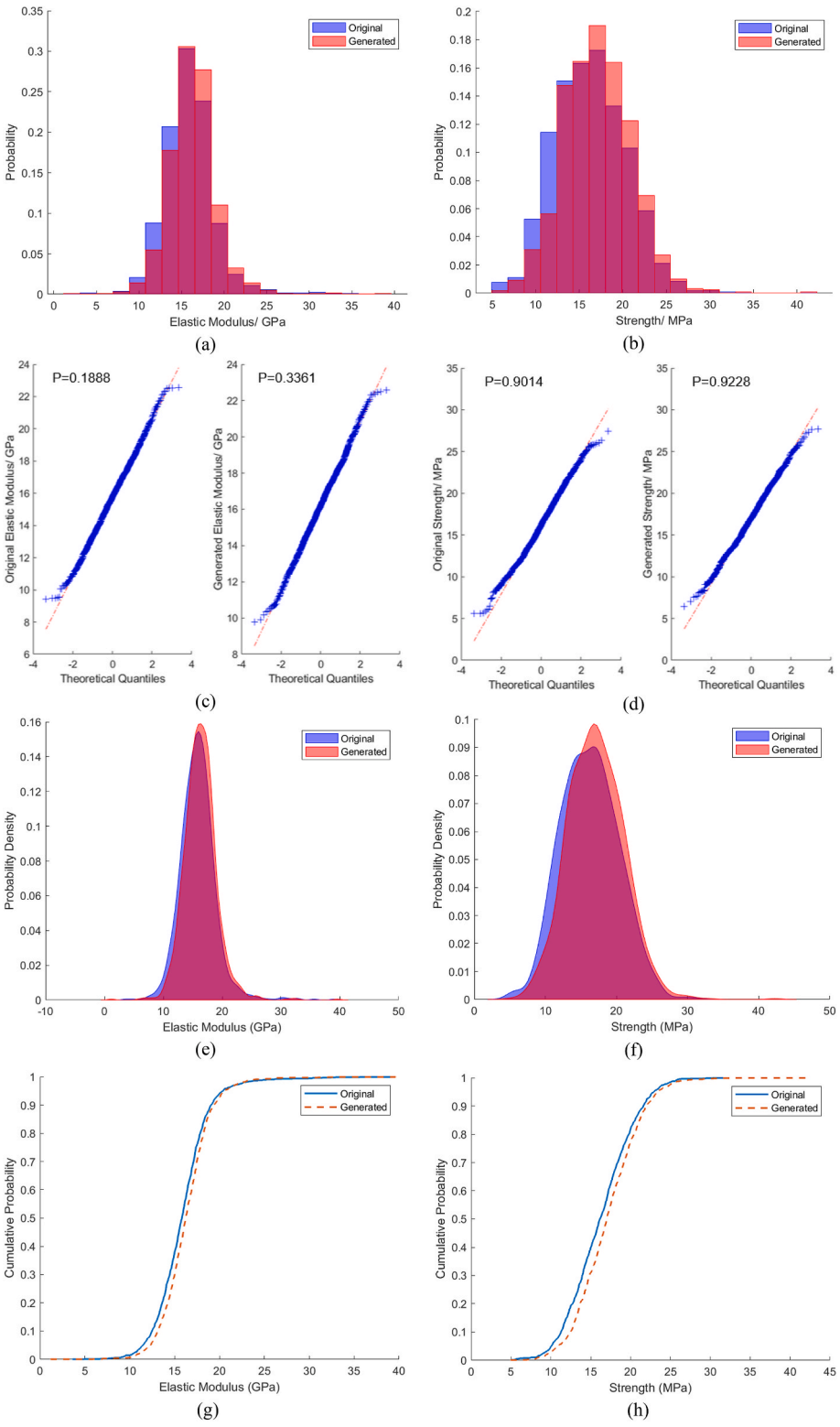

Fig. 15. Statistics of the elastic modulus and strength (a, c, e, g are the histogram, QQ plot, PDF, and CDF of the elastic modulus respectively; b, d, f, h are the histogram, QQ plot, PDF, and CDF of the uniaxial tensile strength respectively.).

- The DDPM of this paper is an unconditional model, i.e., the input of this model is only ordinary Portland cement paste with a single w/c ratio of 0.40. For more general applications, conditional DDPM should be built by incorporating more parameters in the training samples, such as w/c ratios, compositions of clinkers, ages, supplementary cementitious materials, etc. It should be noted that incorporating these parameters does not require significant changes of the DDPM methods and network architectures. These parameters can be encoded and embedded in the input images as the time-embedding procedures described in section 2.3 (Lyu and Ren, 2024).
- The DDPM in this paper only generates two-dimensional images, which of course is a simplification of the real microstructures. Such

limitation is due to the fact that BSE images are two dimensional. If sufficient 3D images are available (such as from XCT), then the DDPM in this paper can be directly adjusted to use 3D volumes as input by simply replacing the 2D convolutional layer as the 3D convolutional layer. If only a limited number of 3D images are available, then a 3D reconstruction model needs to be built to map the 2D images from the DDPM to 3D volumes. This 3D reconstruction model can also use the DDPM, which often uses a few 2D slices as input and generates 3D output (Hui et al., 2024).

- In this paper, the validation of the generated images is mainly by micromechanical analysis. Broader validation schemes are recommended to also include transport properties (such as water permeability, chloride diffusion), volume stability, etc.
- This paper shows that DDPM can extract the latent representation of the microstructure of cement paste. Such latent representation can be used to either predict the material properties or optimize the material design. In (Zheng et al., 2023b), VAE was used as both the image generator, and its low-dimensional latent variables were used to predict the material properties and conduct the material optimization. For DDPM, despite its advantages in image quality, the extraction of latent representation is not straightforward. Recent studies (Yang and Wang, 2023) show that the noise predicted at a few steps are useful for such representation learning, which then are useful for property predictions. Following this approach, a conditional DDPM-based inverse design tool for cementitious materials can also be constructed.

5. Conclusions

This study introduces the application of Denoising Diffusion Probabilistic Models (DDPM) for the synthesis of high-fidelity cement paste microstructures. Comparing with conventional approaches such as SEM, XCT, and other microstructure generation methods, the DDPM approach offers significant improvements in efficiency and cost-effectiveness, while also achieving stable training and high-fidelity image generation without the limitation of phase assumption.

The synthesized microstructures exhibit physically meaningful distributions of LD-CSH, HD-CSH, and UHC, aligning well with classical theories of Portland cement hydration. The generated images maintain nearly identical greyscale statistics and phase assemblage to the original microstructures, ensuring accurate prediction of micromechanical properties such as elastic modulus and uniaxial tensile strength. Additionally, phase connectivity analysis via the lineal path function confirms that the model successfully captures spatial relationships within the microstructures, a critical aspect often overlooked in prior studies.

The established DDPM framework provides a robust foundation for multiscale modeling of cementitious materials, serving as a high-fidelity input for property and behavior analysis. The ability to generate realistic microstructures strengthens the link between microstructure and material properties, facilitating deeper insights into cementitious materials.

Future work should focus on expanding the model's applicability by incorporating diverse parameters such as water/cement ratio, curing age, and supplementary cementitious materials. Integrating broader validation frameworks, including transport properties and volume stability, will further enhance its practical value. Additionally, leveraging the latent representations learned by DDPM for material property predictions and optimization can drive innovative design and analysis of cementitious materials.

CRediT authorship contribution statement

Minfei Liang: Writing – review & editing, Writing – original draft, Software, Investigation, Data curation, Conceptualization. **Kun Feng:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Shan He:** Writing – review & editing, Investigation, Data curation. **Yidong Gan:** Writing – review & editing, Formal analysis. **Yu Zhang:** Writing – review & editing, Formal analysis. **Erik Schlangen:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Branko Šavija:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Branko Šavija acknowledges the financial support of the European Research Council (ERC) within the framework of the ERC Starting Grant Project "Auxetic Cementitious Composites by 3D printing (ACC-3D)", Grant Agreement Number 101041342. Kun Feng acknowledges the financial support of the National Key R&D Program of China, Grant Number 2021YFB2600900, and the National Natural Science Foundation of China, Grant Number 52378418. Yidong Gan acknowledges the financial support of the National Natural Science Foundation of China, Grant Number 52408261.

Data availability

Data will be made available on request.

References

Arjovsky, M., Bottou, L., 2017. Towards Principled Methods for Training Generative Adversarial Networks.

Azqadan, E., Jahed, H., Arami, A., 2023. Predictive microstructure image generation using denoising diffusion probabilistic models. Acta Mater. 261, 119406. https://doi.org/10.1016/j.actamat.2023.119406.

Bentz, D.P., 1997. Three-dimensional computer simulation of Portland cement hydration and microstructure development. J. Am. Ceram. Soc. 80, 3–21. https://doi.org/10.1111/j.1151-2916.1997.tb02785.x.

Bentz, D.P., 1999. Modelling cement microstructure: pixels, particles, and property prediction. Mater. Struct. 32, 187–195. https://doi.org/10.1007/BF02481514.

Bernard, O., Ulm, F.-J., Lemarchand, E., 2003. A multiscale micromechanics-hydration model for the early-age elastic properties of cement-based materials. Cem Concr Res 33, 1293–1309. https://doi.org/10.1016/S0008-8846(03)00039-5.

Bishnoi, S., Scrivener, K.L., 2009. μic: a new platform for modelling the hydration of cements. Cem Concr Res 39, 266–274. https://doi.org/10.1016/j.cemconres.2008.12.002.

Chang, Z., Zhang, H., Schlangen, E., Šavija, B., 2020. Lattice fracture model for concrete fracture Revisited: calibration and validation. Appl. Sci. 10, 4822. https://doi.org/10.3390/app10144822.

Chang, Z., Zhang, H., Liang, M., Schlangen, E., Šavija, B., 2022. Numerical simulation of elastic buckling in 3D concrete printing using the lattice model with geometric nonlinearity. Autom Constr 142, 104485. https://doi.org/10.1016/j.autcon.2022.104485.

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. Sci. Rep. 10, 13307. https://doi.org/10.1038/s41598-020-70149-0.

Chung, S.-Y., Han, T.-S., Kim, S.-Y., Lee, T.-H., 2014. Investigation of the permeability of porous concrete reconstructed using probabilistic description methods. Constr. Build. Mater. 66, 760–770. https://doi.org/10.1016/j.conbuildmat.2014.06.013.

G. Constantinides, F.-J. Ulm, The Elastic Properties of Calcium-Leached Cement Pastes and Mortars: A Multi-Scale Investigation, n.d.

Constantinides, G., Ulm, F.J., 2004. The effect of two types of C-S-H on the elasticity of cement-based materials: results from nanoindentation and micromechanical modeling. Cem Concr Res 34, 67–80. https://doi.org/10.1016/S0008-8846(03)00230-8.

Elfwing, S., Uchibe, E., Doya, K., 2017. Sigmoid-weighted linear units for neural network function approximation. In: Reinforcement Learning.

Gan, Y., Vandamme, M., Zhang, H., Chen, Y., Schlangen, E., van Breugel, K., Šavija, B., 2020. Micro-cantilever testing on the short-term creep behaviour of cement paste at micro-scale. Cem Concr Res 134. https://doi.org/10.1016/j.cemconres.2020.106105.

Gan, Y., Romero Rodriguez, C., Zhang, H., Schlangen, E., van Breugel, K., Šavija, B., 2021. Modeling of microstructural effects on the creep of hardened cement paste using an experimentally informed lattice model. Comput. Aided Civ. Infrastruct. Eng. 36, 560–576. https://doi.org/10.1111/mice.12659.

Gao, X., Wei, Y., Huang, W., 2017. Effect of individual phases on multiscale modeling mechanical properties of hardened cement paste. Constr. Build. Mater. 153, 25–35. https://doi.org/10.1016/j.conbuildmat.2017.07.074.
Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative Adversarial Networks.
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.
Gostick, J., Khan, Z., Tranter, T., Kok, M., Agnaou, M., Sadeghi, M., Jervis, R., 2019. PoreSpy: a Python toolkit for quantitative analysis of porous media images. J. Open Source Softw. 4, 1296. https://doi.org/10.21105/joss.01296.
Han, T.-S., Zhang, X., Kim, J.-S., Chung, S.-Y., Lim, J.-H., Linder, C., 2018. Area of lineal-path function for describing the pore microstructures of cement paste and their relations to the mechanical properties simulated from μ-CT microstructures. Cem. Concr. Compos. 89, 1–17. https://doi.org/10.1016/j.cemconcomp.2018.02.008.
Han, T.-S., Eum, D., Kim, S.-Y., Kim, J.-S., Lim, J.-H., Park, K., Stephan, D., 2023. Multi-scale analysis framework for predicting tensile strength of cement paste by combining experiments and simulations. Cem. Concr. Compos. 139, 105006. https://doi.org/10.1016/j.cemconcomp.2023.105006.
He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition.
Ho, J., Jain, A., Abbeel, P., 2020. Denoising Diffusion Probabilistic Models.
Hong, S.-W., Kim, S.-Y., Park, K., Terada, K., Lee, H., Han, T.-S., 2024. Mechanical property evaluation of 3D multi-phase cement paste microstructures reconstructed using generative adversarial networks. Cem. Concr. Compos. 152, 105646. https://doi.org/10.1016/j.cemconcomp.2024.105646.
Hu, C., Li, Z., 2014. Micromechanical investigation of Portland cement paste. Constr. Build. Mater. 71, 44–52. https://doi.org/10.1016/j.conbuildmat.2014.08.017.
Hu, C., Li, Z., 2015. A review on the mechanical properties of cement-based materials measured by nanoindentation. Constr. Build. Mater. 90, 80–90. https://doi.org/10.1016/j.conbuildmat.2015.05.008.
Hu, C., Gao, Y., Zhang, Y., Li, Z., 2016. Statistical nanoindentation technique in application to hardened cement pastes: influences of material microstructure and analysis method. Constr. Build. Mater. 113, 306–316. https://doi.org/10.1016/j.conbuildmat.2016.03.064.
Hu, C., Hou, D., Li, Z., 2017. Micro-mechanical properties of calcium sulfoaluminate cement and the correlation with microstructures. Cem. Concr. Compos. 80, 10–16. https://doi.org/10.1016/J.CEMCONCOMP.2017.02.005.
Hui, M., Wei, Z., Zhu, H., Xia, F., Zhou, Y., 2024. MicroDiffusion: Implicit Representation-Guided Diffusion for 3D Reconstruction from Limited 2D Microscopy Projections.
Kim, S.-Y., Kim, J.-S., Kang, J.W., Han, T.-S., 2019. Construction of virtual interfacial transition zone (ITZ) samples of hydrated cement paste using extended stochastic optimization. Cem. Concr. Compos. 102, 84–93. https://doi.org/10.1016/j.cemconcomp.2019.04.012.
Kim, Y., Park, H.K., Jung, J., Asghari-Rad, P., Lee, S., Kim, J.Y., Jung, H.G., Kim, H.S., 2021. Exploration of optimal microstructure and mechanical properties in continuous microstructure space using a variational autoencoder. Mater. Des. 202, 109544. https://doi.org/10.1016/j.matdes.2021.109544.
Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization.
Kingma, D.P., Welling, M., 2013. Auto-Encoding Variational Bayes.
Liang, S., Wei, Y., Wu, Z., 2017. Multiscale modeling elastic properties of cement-based materials considering imperfect interface effect. Constr. Build. Mater. 154, 567–579. https://doi.org/10.1016/j.conbuildmat.2017.07.196.
Liang, M., Gan, Y., Chang, Z., Wan, Z., Schlangen, E., Šavija, B., 2022. Microstructure-informed deep convolutional neural network for predicting short-term creep modulus of cement paste. Cem Concr Res 152, 106681. https://doi.org/10.1016/J.CEMCONRES.2021.106681.
Liang, M., He, S., Gan, Y., Zhang, H., Chang, Z., Schlangen, E., Šavija, B., 2023a. Predicting micromechanical properties of cement paste from backscattered electron (BSE) images by computer vision. Mater. Des. 229, 111905. https://doi.org/10.1016/j.matdes.2023.111905.
Liang, M., Zhang, Y., He, S., Chen, Y., Schlangen, E., Šavija, B., 2023b. On the chemo-mechanical evolution process of high-volume slag cement paste. Constr. Build. Mater. 400, 132891. https://doi.org/10.1016/j.conbuildmat.2023.132891.
Lyu, X., Ren, X., 2024. Microstructure reconstruction of 2D/3D random materials via diffusion-based deep generative models. Sci. Rep. 14, 5041. https://doi.org/10.1038/s41598-024-54861-9.
Metz, L., Poole, B., Pfau, D., Sohl-Dickstein, J., 2016. Unrolled Generative Adversarial Networks.
Murgas, B., Stickel, J., Ghosh, S., 2024. Generative adversarial network (GAN) enabled Statistically equivalent virtual microstructures (SEVM) for modeling cold spray formed bimodal polycrystals. npj Comput. Mater. 10, 32. https://doi.org/10.1038/s41524-024-01219-4.
Prince, Simon J.D., 2023. Understanding Deep Learning. The MIT Press.
Qian, Zhiwei, 2012. Modeling of Fracture Processes in Cementitious Materials. Delft University of Technology.
Rhardane, A., Grondin, F., Alam, S.Y., 2020. Development of a micro-mechanical model for the determination of damage properties of cement pastes. Constr. Build. Mater. 261, 120514. https://doi.org/10.1016/j.conbuildmat.2020.120514.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional Networks for Biomedical Image Segmentation.
Scrivener, K., Snellings, R., Lothenbach, B., Press, C.R.C.(eds.), 2016. A practical guide to microstructural analysis of cementitious materials, 540. Crc Press, Boca Raton, FL, USA.
Šmilauer, V., Bittnar, Z., 2006. Microstructure-based micromechanical prediction of elastic properties in hydrating cement paste. Cem Concr Res 36, 1708–1718. https://doi.org/10.1016/j.cemconres.2006.05.014.
Safiuddin, M., Reddy, C.L., Vasantada, G., Harsha, C., Gangolu, S., 2021. Establishing Process-Structure Linkages Using Generative Adversarial Networks.
Sahimi, M., Tahmasebi, P., 2021. Reconstruction, optimization, and design of heterogeneous materials and media: basic principles, computational algorithms, and applications. Phys. Rep. 939, 1–82. https://doi.org/10.1016/j.physrep.2021.09.003.
Šavija, B., Pacheco, J., Schlangen, E., 2013a. Lattice modeling of chloride diffusion in sound and cracked concrete. Cem. Concr. Compos. 42, 30–40. https://doi.org/10.1016/j.cemconcomp.2013.05.003.
Šavija, B., Luković, M., Pacheco, J., Schlangen, E., 2013b. Cracking of the concrete cover due to reinforcement corrosion: a two-dimensional lattice model study. Constr. Build. Mater. 44, 626–638. https://doi.org/10.1016/j.conbuildmat.2013.03.063.
Šavija, B., Luković, M., Schlangen, E., 2014. Lattice modeling of rapid chloride migration in concrete. Cem Concr Res 61–62, 49–63. https://doi.org/10.1016/j.cemconres.2014.04.004.
Šavija, B., Zhang, H., Schlangen, E., 2020. Micromechanical testing and modelling of blast furnace slag cement pastes. Constr. Build. Mater. 239. https://doi.org/10.1016/j.conbuildmat.2019.117841.
Schlangen, E., Garboczi, E.J., 1997a. Fracture simulations of concrete using lattice models: computational aspects. Eng. Fract. Mech. 57, 319–332. https://doi.org/10.1016/S0013-7944(97)00010-6.
Schlangen, E., Garboczi, E.J., 1997b. FRACTURE SIMULATIONS OF CONCRETE USING LATTICE MODELS: COMPUTATIONAL ASPECTS.
Schlangen, E., van Mier, J.G.M., 1992. Simple lattice model for numerical simulation of fracture of concrete materials and structures. Mater. Struct. 25, 534–542. https://doi.org/10.1007/BF02472449.
Scrivener, K.L., 2004. Backscattered electron imaging of cementitious microstructures: understanding and quantification. Cem. Concr. Compos. 26, 935–945. https://doi.org/10.1016/j.cemconcomp.2004.02.029.
Taylor, H.F.W., 1997. Cement Chemistry. Thomas Telford Publishing. https://doi.org/10.1680/cc.25929.
Tennis, P.D., Jennings, H.M., 2000. A model for two types of calcium silicate hydrate in the microstructure of Portland cement pastes. Cem Concr Res 30, 855–863. https://doi.org/10.1016/S0008-8846(00)00257-X.
Torquato, S., 2002a. Random Heterogeneous Materials. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4757-6355-3.
Torquato, S., 2002b. Random Heterogeneous Materials. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4757-6355-3.
Ulm, F.J., Vandamme, M., Jennings, H.M., Vanzo, J., Bentivegna, M., Krakowiak, K.J., Constantinides, G., Bobko, C.P., Van Vliet, K.J., 2010. Does microstructure matter for statistical nanoindentation techniques? Cem. Concr. Compos. 32, 92–99. https://doi.org/10.1016/J.CEMCONCOMP.2009.08.007.
Vahdat, A., Kautz, J., 2020. NVAE: A Deep Hierarchical Variational Autoencoder.
Van Breugel, K., 1993. Simulation of Hydration and Formation of Structure in Hardening Cement-Based Materials. Delft University of Technology.
van Breugel, K., 1995. Numerical simulation of hydration and microstructural development in hardening cement-based materials (I) theory. Cem Concr Res 25, 319–331. https://doi.org/10.1016/0008-8846(95)00017-8.
Vandamme, M., 2002. The Nanogranular Origin of Concrete Creep: A Nanoindentation Investigation of Microstructure and Fundamental Properties of Calcium-Silicate-Hydrates. Massachusetts Institute of Technology.
Vandamme, M., Ulm, F.-J., 2013. Nanoindentation investigation of creep properties of calcium silicate hydrates. Cem Concr Res 52, 38–52. https://doi.org/10.1016/j.cemconres.2013.05.006.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is All You Need.
Wong, H.S., Head, M.K., Buenfeld, N.R., 2006. Pore segmentation of cement-based materials from backscattered electron images. Cem Concr Res 36, 1083–1090. https://doi.org/10.1016/j.cemconres.2005.10.006.
Wu, Y., He, K., 2018. Group Normalization.
Yang, X., Wang, X., 2023. Diffusion Model as Representation Learner.
Ye, G., van Breugel, K., Fraaij, A.L.A., 2003. Three-dimensional microstructure analysis of numerically simulated cementitious materials. Cem Concr Res 33, 215–222. https://doi.org/10.1016/S0008-8846(02)00889-X.
Zhang, J., Scherer, G.W., 2011. Comparison of methods for arresting hydration of cement. Cem Concr Res 41, 1024–1036. https://doi.org/10.1016/J.CEMCONRES.2011.06.003.
Zhang, H., Šavija, B., Figueiredo, S.C., Schlangen, E., 2017. Experimentally validated multi-scale modelling scheme of deformation and fracture of cement paste. Cem Concr Res 102, 175–186. https://doi.org/10.1016/J.CEMCONRES.2017.09.011.
Zhang, H., Šavija, B., Xu, Y., Schlangen, E., 2018a. Size effect on splitting strength of hardened cement paste: experimental and numerical study. Cem. Concr. Compos. 94, 264–276. https://doi.org/10.1016/J.CEMCONCOMP.2018.09.018.
Zhang, H., Šavija, B., Schlangen, E., 2018b. Towards understanding stochastic fracture performance of cement paste at micro length scale based on numerical simulation. Constr. Build. Mater. 183, 189–201. https://doi.org/10.1016/j.conbuildmat.2018.06.167.
Zhang, H., Šavija, B., Luković, M., Schlangen, E., 2019a. Experimentally informed micromechanical modelling of cement paste: an approach coupling X-ray computed tomography and statistical nanoindentation. Compos. B Eng. 157, 109–122. https://doi.org/10.1016/J.COMPOSITESB.2018.08.102.
Zhang, H., Xu, Y., Gan, Y., Chang, Z., Schlangen, E., Šavija, B., 2019b. Combined experimental and numerical study of uniaxial compression failure of hardened cement paste at micrometre length scale. Cem Concr Res 126, 105925. https://doi.org/10.1016/J.CEMCONRES.2019.105925.
Zhang, H., Xu, Y., Gan, Y., Chang, Z., Schlangen, E., Šavija, B., 2020. Microstructure informed micromechanical modelling of hydrated cement paste: techniques and

challenges. Constr. Build. Mater. 251, 118983. https://doi.org/10.1016/J.CONBUILDMAT.2020.118983.
Zhao, X., Wang, L., Li, Q., Chen, H., Liu, S., Hou, P., Ye, J., Pei, Y., Wu, X., Yuan, J., Gao, H., Yang, B., 2025. 3D microstructural generation from 2D images of cement paste using generative adversarial networks. Cem Concr Res 187, 107726. https://doi.org/10.1016/j.cemconres.2024.107726.
Zheng, L., Karapiperis, K., Kumar, S., Kochmann, D.M., 2023a. Unifying the design space and optimizing linear and nonlinear truss metamaterials by generative modeling. Nat. Commun. 14, 7563. https://doi.org/10.1038/s41467-023-42068-x.
Zheng, L., Karapiperis, K., Kumar, S., Kochmann, D.M., 2023b. Unifying the design space and optimizing linear and nonlinear truss metamaterials by generative modeling. Nat. Commun. 14, 7563. https://doi.org/10.1038/s41467-023-42068-x.