Statistical pre-detection of cracks in 3D concrete images for computational efficiency

Vitalii Makogin[a,*], Duc Nguyen[a,**], Evgeny Spodarev[a]

[a]*Institute of Stochastics, Ulm University, Helmholtzstraße 18, Ulm, 89075, Baden-Württemberg, Germany*

Abstract

In practical applications, effectively segmenting cracks in large-scale computed tomography (CT) images holds significant importance for understanding the structural integrity of materials. However, classical methods and Machine Learning algorithms often incur high computational costs when dealing with the substantial size of input images. Hence, a robust algorithm is needed to pre-detect crack regions, enabling focused analysis and reducing computational overhead. The proposed approach addresses this challenge by offering a streamlined method for identifying crack regions in CT images with high probability. By efficiently identifying areas of interest, our algorithm allows for a more focused examination of potential anomalies within the material structure. Through comprehensive testing on both semi-synthetic and real 3D CT images, we validate the efficiency of our approach in enhancing crack segmentation while reducing computational resource requirements.

Keywords:
CUSUM, multiple hypotheses testing, Hessian-based filter, crack detection, classification.

1. Introduction

Concrete serves as the foundational material for various structures, including buildings and bridges, underscoring the importance of ensuring its quality, durability, and mechanical stability. To gain a better understanding of the mechanisms and

*Dedicated to the memory of Dr. Vitalii Makogin (12.12.1987 - 08.05.2024)

**Corresponding author

Email addresses: tran-1.nguyen@uni-ulm.de (Duc Nguyen), evgeny.spodarev@uni-ulm.de (Evgeny Spodarev)

causes of cracks, it is essential to subject concrete specimens to stress tests, which is valuable to prevent the occurrence of emergency conditions. For instance, the damage and fracture evolution in concrete were studied in [19] under cyclic loading. An overview study about this type of material was also introduced in [18].

Nevertheless, analyzing the CT image of concrete samples after a stress test can facilitate the study of crack propagation. Typically, a 3D CT image comprises two components: cracks and background. Cracks exhibit lower gray values compared to the background. This characteristic suggests the use of segmentation methods, involving the application of a global threshold or a pixel/voxel-wise analysis. Various methods for image segmentation and crack detection fall into distinct categories, including thresholding segmentation based on assumptions about the distribution of gray values [35, 28], edge-detection techniques [13, 1] that consider local structures, and region-growing approaches like percolation-based methods [33]. However, dealing with 3-dimensional images introduces significant challenges due to the high noise levels and the intricate nature of crack structures. Detecting cracks in 3D images introduces challenges due to computational costs and a higher risk of misclassification. Filters like the Frangi filter and Sheet filter [17, 32] prove to be effective in identifying flat-like or vessel-like structures within the material under appropriate settings. Leveraging the results from the Frangi filter or Sheet filter, the Hessian-based percolation method [15] is then applied, enhancing crack detection performance. In a study on semi-synthetic images [5], it was demonstrated that, among classical methods, the Frangi filter, Sheet filter, and Hessian-based percolation filter yield optimal results when their parameters are tuned appropriately.

Obviously, to tackle such challenges, machine learning and deep learning algorithms have been recently considered. Methodologies such as convolutional neural networks, random forests, 3D-Unet can be applied for crack segmentation. However, with the lack of training data, the accuracy of those methods might not be sufficient. As an alternative, semi-synthetic data can be generated (based on minimum-weight surfaces in bounded Voronoi diagrams [24, 26]), exhibiting similar geometry to real-world input. However, when dealing with extremely large input images (e.g., $10000^2 \times 2000$) which modern scanners produce, efficient storage and computational power become crucial. Hence, it is necessary to pre-localize the cracks to reduce the size of the input data to be used in artificial intelligence methods. This issue imposes a natural restriction to use only methods with linear complexity. For that, it is advisable to partition the input image into smaller subregions, subsequently inspecting each region in this partition for cracks. The ultimate crack detection may then by performed by a deep learning model proposed in [6]. Based on the partition, we construct a realization of a geometric random field such that cracks in the input image are treated

as anomaly regions within this field.

The change-point problem for stationary random fields $\{\xi_k, k \in W\}, W \subset \mathbb{Z}^n$, as discussed in [11], involves defining changes in parameters such as mean, variance, spectral density, etc., within $\{\xi_k, k \in W\}, W \subset \mathbb{Z}^n$. The Cumulative Sum (CUSUM) statistic is a crucial tool for detecting clusters in this context. As anomaly detection, the change-point analysis has been extended in [2, 3] to 3D CT images to identify changes in the mean and entropy of the local directional distribution of fibers. Assuming the random field $\{\xi_k, k \in W\}, W \subset \mathbb{Z}^n$ to be stationary and m-dependent, the tail probability of CUSUM statistics is upper-bounded [14], enabling to control the Type-I errors during scan statistic tests.

An alternative strategy involves simultaneous testing within all scanning windows, known as multiple hypotheses testing. Early approaches, such as the Bonferroni correction [10] and the Šidák correction [29], aimed to rigorously control the Family-Wise Error Rate (FWER) to prevent inflation of the probability of making at least one false discovery. Subsequent methods, introduced by [21, 22, 30], included step-down and step-up procedures that dynamically adapt based on observed p-values, striking a balance between error control and sensitivity. In large-scale hypothesis testing, rather than maintaining an extremely conservative test, one can relax the rejection rule, allowing for more false positives and increased power. This led to the introduction of the False Discovery Rate (FDR) concept by [7], which also proposed a testing procedure resulting in a significant gain in power.

In this study, we have developed a comprehensive framework comprising three main phases to improve the detection of cracks in 3D CT concrete images. The initial phase involves pre-processing the 3D CT concrete image to distinguish cracks from the surrounding material. Following basic crack segmentation, we study the geometry properties of each subregion within a partition, offering a preliminary indication of crack presence.

After that, we construct a multivariate random field observed within a collection of parameterized scanning windows, which contains the information gathered in the previous phase. The final phase involves applying the Benjamini-Hochberg multiple testing procedure. This step classifies anomaly regions within the concrete, providing a refined understanding of potential areas of interest. The whole procedure is described in Diagram 1.

The paper is organized as follows: In Section 2, we present various classical edge-detection methods and the implementation of the Hessian-based filter on input images. A comparison between these methods over semi-synthetic images is also discussed. Furthermore, the selection and computation of a geometric multivariate random field $\{\xi_k, k \in W \subset \mathbb{R}^n\}, W \subset \mathbb{Z}^n$ is motivated, providing a suitable framework

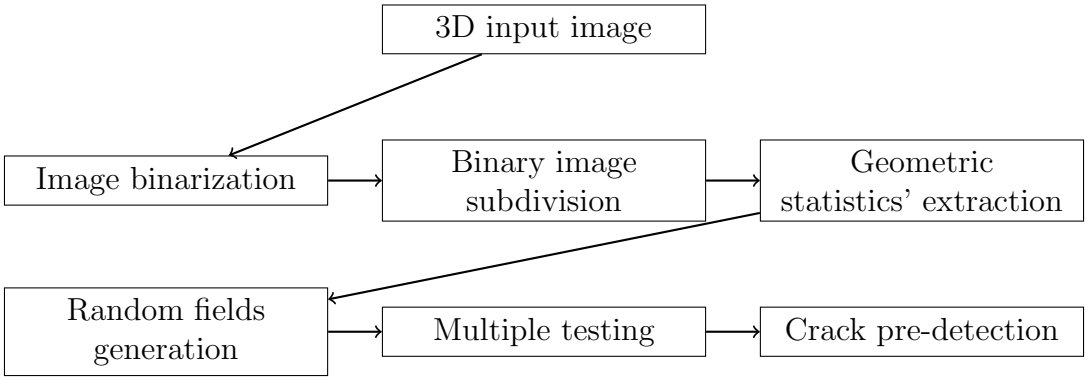


Figure 1: 6-step processing pipeline diagram for crack localization in 3D concrete images

for hypothesis testing. In Section 3, we use CUSUM statistic test in the context of multiple hypotheses tests in order to find anomaly regions within a realization of $\{\xi_k, k \in W \subset \mathbb{R}^n\}, W \subset \mathbb{Z}^n$. Section 4 presents numerical results for both semi-synthetic and real images, with a discussion on the advantages and disadvantages of the entire framework. Finally, in Section 5, we conclude the paper, summarizing key findings and suggesting potential challenges for future research.

2. Crack segmentation

In this section, we present several classical methods proposed in [17, 32] including Frangi filter and Sheet filter. The Hessian-based percolation [15, 33] is used for crack tracking once candidate voxels for a crack are identified. The common idea for these methods is to locally compute the Hessian matrix of an image, which captures its second-order partial derivatives, providing a quantitative representation of local image structures and intensity variations based on its eigenvalues. This approach possesses both pros and contras, notably in terms of computational expenses. Therefore, a simpler method will be introduced to overcome this crucial difficulty. Let W be a subset of $\mathbb{Z}^3$. The grayscale image $I = \{I(p) \in [0,1], p \in W\}$ is employed to represent the input data.

2.1. Classical crack segmentation methods

The smoothed second-order partial derivatives of the input image I are computed as follows:

$$\frac{\tilde{\partial}^2 I}{\partial p_i \partial p_j}(p,\sigma) = \sigma I(p) * \frac{\partial^2}{\partial p_i \partial p_j} G(p,\sigma),\ i,j = 1,2,3,\ p = (p_1,p_2,p_3)^T \in W,$$

where G is the 3-dimensional Gaussian kernel $G(p,\sigma) = (2\pi\sigma^2)^{-3/2}\exp\{-\|p\|_2^2/(2\sigma^2)\}$ with scale parameter $\sigma > 0$ and $\|\cdot\|_2$ the Euclidean norm in $\mathbb{R}^3$. Here $*$ denotes the usual convolution operation. The Hessian matrix $H(p,\sigma)$ of an image I at a voxel $p \in W$ is given by

$$H(p,\sigma) = \begin{bmatrix} \frac{\tilde{\partial}^2 I}{\partial p_1^2} & \frac{\tilde{\partial}^2 I}{\partial p_1 \partial p_2} & \frac{\tilde{\partial}^2 I}{\partial p_1 \partial p_3} \\ \frac{\tilde{\partial}^2 I}{\partial p_2 \partial p_1} & \frac{\tilde{\partial}^2 I}{\partial p_2^2} & \frac{\tilde{\partial}^2 I}{\partial p_2 \partial p_3} \\ \frac{\tilde{\partial}^2 I}{\partial p_3 \partial p_1} & \frac{\tilde{\partial}^2 I}{\partial p_3 \partial p_2} & \frac{\tilde{\partial}^2 I}{\partial p_3^2} \end{bmatrix}$$

The eigenvalues of H play a crucial role in describing and characterizing local structures within an image through considerations such as their magnitudes and ratios. Represented in ascending order as $|\lambda_1(p,\sigma)| \leq |\lambda_2(p,\sigma)| \leq |\lambda_3(p,\sigma)|$, these eigenvalues at voxel p offer crucial insights into the crack structure of concrete samples. The visualization of these eigenvalues is exemplified in Figure 2.

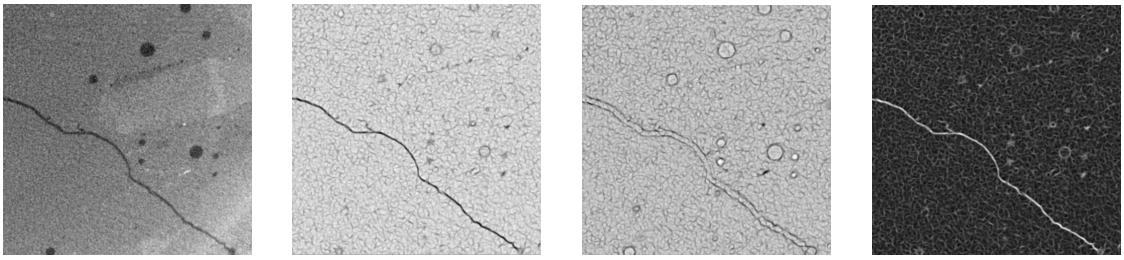


Figure 2: From the left to the right: 2d slices of input image I and three images containing the eigenvalues $\lambda_1(p,\sigma), \lambda_2(p,\sigma), \lambda_3(p,\sigma)$ of H.

The Sheet filter and Frangi filter predominantly rely on the numerical values of $\lambda_1(p,\sigma), \lambda_2(p,\sigma), \lambda_3(p,\sigma)$. They were designed to detect lower dimensional image features such as vessels and cracks. Given the scale parameter σ as well as parameters $\delta > 0, \rho \in (0,1]$ at voxel p, denote by $S(p,\sigma)$ and $F(p,\sigma)$ the outcomes produced by Sheet filter and Frangi filter, respectively:

$$S(p,\sigma) = \begin{cases} \lambda_3(p,\sigma) \cdot g(\lambda_1(p,\sigma), \lambda_3(p,\sigma)) \cdot g(\lambda_2(p,\sigma), \lambda_3(p,\sigma)), & \lambda_3(p,\sigma) > 0, \\ 0, & \text{else}, \end{cases}$$

where

$$g(\lambda_s(p,\sigma),\lambda_t(p,\sigma)) = \begin{cases} \left[1+\frac{\lambda_s(p,\sigma)}{|\lambda_t(p,\sigma)|}\right]^{\delta}, & \lambda_s(p,\sigma)\le 0, |\lambda_t(p,\sigma)|\ge|\lambda_s(p,\sigma)|, \\ \left[1-\rho\frac{\lambda_s(p,\sigma)}{|\lambda_t(p,\sigma)|}\right]^{\delta}, & \lambda_s(p,\sigma)>0, |\lambda_t(p,\sigma)|\ge\rho\lambda_s(p,\sigma), \\ 0, & \text{else,} \end{cases}$$

and

$$F(p,\sigma) = \begin{cases} \exp\left(-\frac{Q_A^2(p)}{a}\right)\exp\left(-\frac{Q_B^2(p)}{b}\right)\left(1-\exp\left(-\frac{K^2(p)}{c}\right)\right), & \lambda_3(p,\sigma)>0, \\ & \lambda_2(p,\sigma)\neq 0, \\ \exp\left(-\frac{Q_A^2(p)}{a}\right)\left(1-\exp\left(-\frac{K^2(p)}{c}\right)\right), & \lambda_3(p,\sigma)>0, \\ & \lambda_2(p,\sigma)=0, \\ 0, & \text{else,} \end{cases}$$

where

$$Q_A(p) := Q_A(p,\sigma) = \left|\frac{\lambda_2(p,\sigma)}{\lambda_3(p,\sigma)}\right|, \quad Q_B(p) := Q_B(p,\sigma) = \frac{|\lambda_1(p,\sigma)|}{\sqrt{|\lambda_2(p,\sigma)\lambda_3(p,\sigma)|}}.$$

and

$$K(p) := K(p,\sigma) = \sqrt{\lambda_1^2(p,\sigma)+\lambda_2^2(p,\sigma)+\lambda_3^2(p,\sigma)}.$$

Both filters rely on assessing the magnitude of the difference between eigenvalues of H, with λ_3 highlighting significant changes along specific directions within the image. Intuitively, the variations between these eigenvalues suggest the existence of a direction at p in which gray values of I vary the most compared to others, indicating the presence of a flat structure that p belongs to.

If a voxel p belongs to a crack, it is highly probable that both $S(p,\sigma)$ and $F(p,\sigma)$ will exhibit higher values compared to voxels belonging to homogeneous parts of I. To distinguish cracks from the homogeneous material, an appropriate global threshold must be applied to filtered images $\{S(p,\sigma), p\in W\}$ and $\{F(p,\sigma), p\in W\}$. Here, both adaptive and manual threshold settings are available. For instance, a suitable threshold range is proposed in [5], assuming synthetic input images with a fixed width for cracks.

However, in practice, the anomalies within a real CT image of a material often do not exhibit ideal conditions, such as a constant width of cracks or significantly higher grayscale contrast at the crack edges, which might make the choice of a threshold T obstacle. To overcome this disadvantage, one can consider an adaptive global threshold $T=\mu(I)+3\cdot sd(I)$, where $\mu(I)$ and $sd(I)$ are the sample mean and the sample standard deviation of all gray values within I, let

Then the binary images S_σ and F_σ can be obtained using T as

$$S^*(p,\sigma) = \mathbb{1}\{S(p,\sigma) \geq \mu(S_\sigma) + 3sd(S_\sigma)\},\ p \in W,$$
$$F^*(p,\sigma) = \mathbb{1}\{F(p,\sigma) \geq \mu(F_\sigma) + 3sd(F_\sigma)\},\ p \in W.$$

It is worth noting that the quality of the binary images produced by both methods heavily depends on the scale parameter σ, recommended to be half of the crack width. These approaches demonstrate robust performance in synthetic input images where cracks are well-designed with a constant width. However, in real-world images, the geometry of cracks is unpredictable, affecting the appropriateness of the choice of σ. To tackle this challenge, a multi-scale approach is proposed, which involves computing outcomes from these filters at various scales and selecting the maximum result. Subsequently, a proper threshold is applied. The multi-scale Frangi filter is defined as:

$$F_S(p) = \max_{\sigma_{min} \leq \sigma \leq \sigma_{max}} F(p,\sigma),$$

where σ_{min} and σ_{max} are chosen in advance. Furthermore, with the assistance of the two aforementioned filters, one can proactively identify a set of voxels, denoted as H, displaying characteristics indicative of potential crack regions. The Hessian-based percolation methodology, structured as an iterative process, systematically explores the neighborhood by commencing the search from all voxels within the set K. The exploration criteria are contingent upon grayscale values. Originally developed for 2D images [33], the methodology has been generalized to 3D case [15]. The algorithmic representation is outlined as follows:

1. Select the set of possible crack voxels H
2. Initialize $P = H$.
3. $\forall p \in H$, let $t = I(p) + \epsilon$ for $\epsilon \in \mathbb{R}$. If $q \in W$ is a neighbor of p, and $I(q) < t$, then add q to P.
4. Let t be the new threshold $t = \max(\max_{q \in P} I(q), t) + \epsilon$.
5. Repeat steps 3 and 4 until $P \cap \partial E \neq \emptyset$, where $M \in \mathbb{Z}, M > 0$ chosen in advance, ∂E is the boundary of the window $\{b \in W, \|b - p\|_\infty = M\}$.
6. If the quantity $|P \cap H|/|P| \geq r$, where $|\cdot|$ is the cardinality of corresponding set, and r is prespecified, then label all voxels in P as crack.
7. Denote by $\tau(p)$ the times that p is taken into account during this process. If $\tau(p) \leq 5$, then label p as a voxel belonging to material.

Obviously, the effectiveness of this algorithm is significantly influenced by the careful selection of the set H, as it directly impacts the ratio $|P \cap H|/|P|$. Furthermore, this method affords the capability to include some disconnected voxels that are presumed to be part of a crack.

2.2. Maximal Hessian Entry filter

With the challenge posed by the variability in crack widths, careful consideration is necessary for computational runtime and storage requirements. Images exceeding dimensions of 600^3 present significant computational challenges in both Hessian matrix computation and eigenvalue extraction. In response to this challenge, an alternative methodology is proposed—one that avoids extracting maximal information from the Hessian matrix and instead relies exclusively on its entries.

From Figure 1, cracks in CT images can manifest as dark regions or disruptions in the intensity pattern, presenting local changes at certain voxels. Due to the flat structure of cracks, if a voxel p belongs to a crack, the gray values along a line tangent to the crack surface will exhibit a pattern similar to a concave function, since the gray values of voxels belonging to homogeneous material are higher than the ones in cracks. Therefore, across all Hessian matrix elements, i.e., the second-order derivatives along six different directions, there should exist at least one element that is non-negative. Choosing the maximal nonnegative entry allows us to capture any sudden changes at voxels belonging to cracks. For an image $I = \{I(p) \in [0,1], p \in W \subset \mathbb{Z}^3\}$, with a single value of σ, the Maximal Hessian Entry filter is defined as follows:

$$L_\sigma = \{L_\sigma(p) = \max_{i,j=1,2,3}(H_{i,j}(p,\sigma), 0), p \in W\}.$$

We apply the same 3σ-rule for the global threshold, which results in a binary image:

$$L^*_\sigma = \{L^*_\sigma(p) = \mathbb{1}\{L_\sigma(p) \geq \mu(L_\sigma) + 3sd(L_\sigma)\}, p \in W\}.$$

Let $\sigma \in \mathcal{S}$, then compute

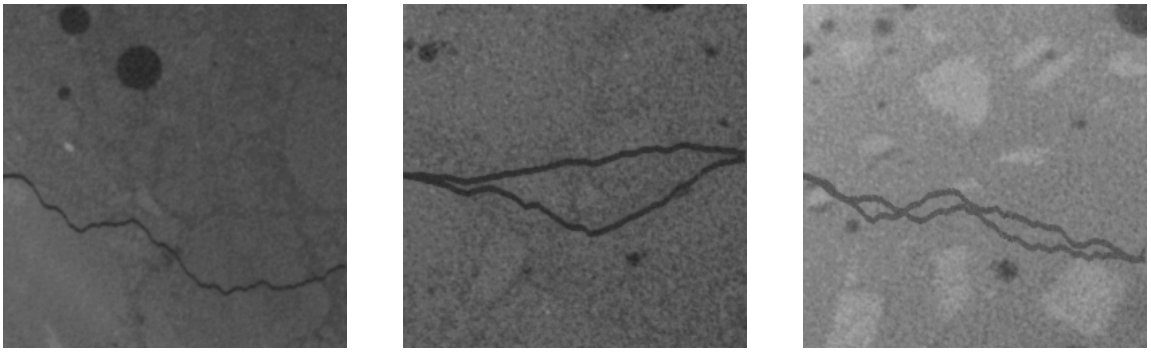
$$L_\mathcal{S} = \{L_\mathcal{S}(p) = \max_{\sigma \in \mathcal{S}} L^*_\sigma(p), p \in W\}.$$

The choice of the discrete set $\mathcal{S}$ can be predetermined according to the nature of the input image. It is important to note that this approach is implemented based solely on the maximal positive entry of the Hessian matrix and a global threshold, leading to a binary image with noticeable noise. However, as observed in the comparison in Section 2.3, the method still preserves the essential structure of cracks, which is beneficial for our subsequent geometric analysis.

2.3. Performance and run time

To compare the performance of the above crack segmentation methods, we executed them on different semi-synthetic images with a resolution of 256^3 and varying crack widths $w = 3; 5$, which are provided by Technical University of Kaiserslautern and

Fraunhofer ITWM, see Figure 3. The algorithm to generate semi-synthetic images and the repository (Zenodo) are provided in [26].

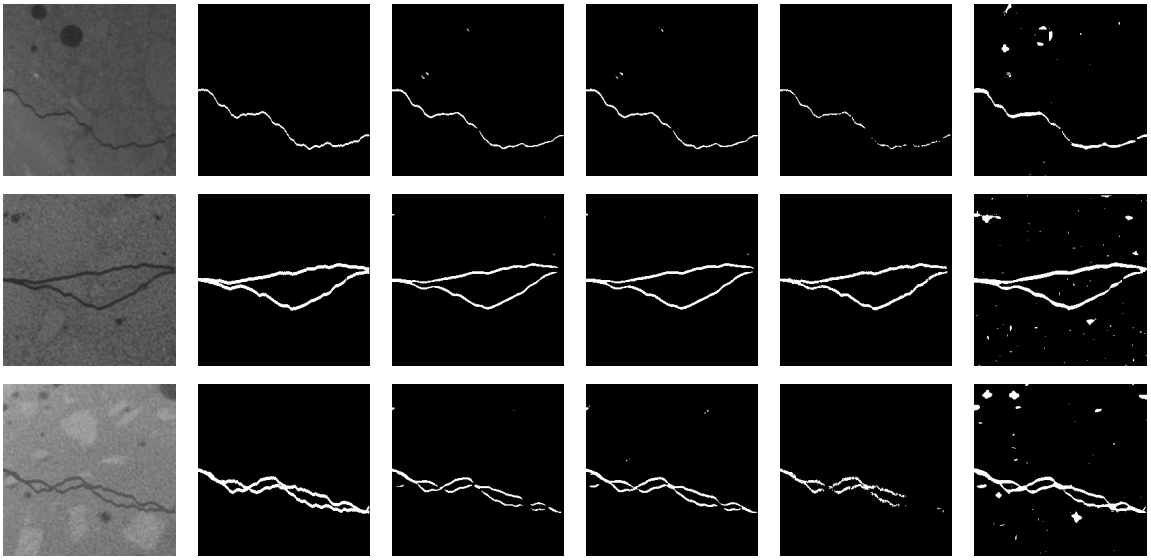


(a) Image I_1 with $w = 3$ (b) Image I_2 with $w = 5$ (c) Image I_3 with $w = 5$

Figure 3: 2D slices of input images with different widths.

The careful selection of these parameters plays a pivotal role in influencing the outcomes of our image processing methods. The comprehensive investigation of optimal parameters for the Frangi filter (FF), Sheet filter (SF), and Hessian percolation (HP) has been conducted in [5] given the Ground truth (GT). For the Frangi filter, when $w = 3$, we selected $\sigma = 1.5, a = 0.3, b = 0.3$, and for $w = 5$, we chose $\sigma = 2.5, a = 0.3, b = 0.3$. For the Sheet filter, when $w = 3$, we set $\sigma = 1.5, \rho = 1, \delta = 1.5$, and for $w = 5$, we used $\sigma = 2.5, \rho = 1, \delta = 1$. For Hessian percolation, when $w = 3$, we specified $\epsilon = -0.5, \tau = 4, r = 0.6, M = 3$, and for $w = 5$, we employed $\epsilon = -0.5, \tau = 5, r = 0.6, M = 5$. For the Hessian Maximal entry filter (HM), with w values of both 3 and 5, we selected the structuring element $\mathcal{S} = \{1, 3, 5\}$.

The results obtained by applying these methods to the input images I_1, I_2, I_3 are showcased in Figure 4. It is evident that all four approaches successfully capture the structural intricacies of cracks within the input images, which highlights the effectiveness of the Hessian Maximal entry filter. However, due to the presence of pores with low gray values in the material, voxels belonging to the edges may be incorrectly identified as cracks. This discrepancy arises from the limited examination of the surrounding structure. Furthermore, by taking the maximum over all binary images, all noise is retained, resulting in a more noisy image compared to the ones from classical filters.



(a) Input image (b) GT (c) FF (d) SF (e) HP (f) HM

Figure 4: From left to right: 2D slices of the input images, ground truth and theirs binary images with $w = 3, 5, 5$, respectively.

To provide a more comprehensive evaluation, the quality of these methods can be assessed through metrics such as precision (P), recall (R), and F1-score (F1), defined as follows:

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN}, \quad F1 = \frac{2PR}{P + R}.$$

Here, TP (true positive) and FP (false positive) represent the numbers of voxels correctly and falsely detected as belonging to a crack, respectively. Similarly, TN (true negative) and FN (false negative) denote the numbers of voxels correctly and falsely detected as material. The performance metrics for these methods are summarized in Table 1. It is easy to see that, from the recall metric, the Maximal Hessian Entry filter is able to capture effectively crack voxels compared with the other methods.

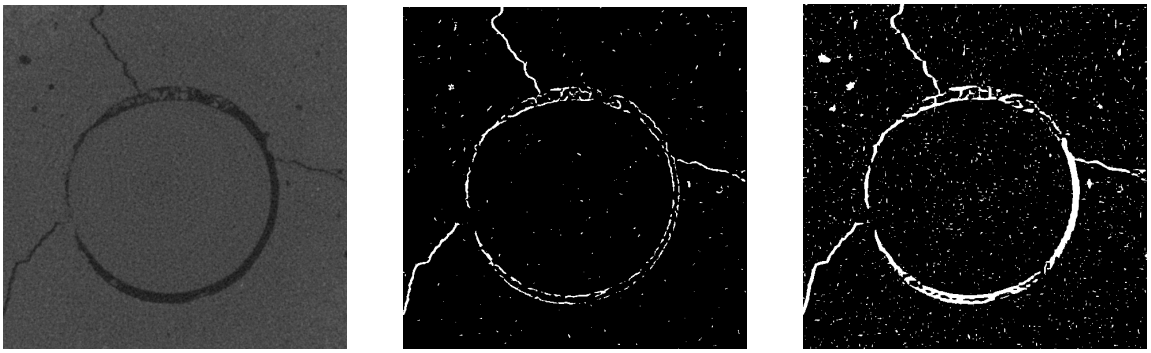
From the setting of the Hessian Maximal entry filter, it becomes apparent that numerous voxels belonging to the material are misclassified as cracks. Consequently, the precision and F1-score of this method are notably lower compared to its counterparts. Despite these limitations, the Hessian Maximal entry filter demonstrates its ability to preserve the structural details of cracks, as indicated by its recall, which is comparable to that of other established methods. The benefits of this approach will be further demonstrated in the applications to real-world images, here we only compare the

		Precision	Recall	F1-score
Image I_1	Frangi filter	0.5751666	0.8662149	0.6913061
	Sheet filter	0.5982854	0.8658113	0.7076066
	Hessian Percolation	0.9013737	0.7901476	0.8421038
	Hessian Maximal entry filter	0.2016482	0.8860226	0.3285275
Image I_2	Frangi filter	0.9879916	0.6631641	0.7936266
	Sheet filter	0.9743654	0.7249849	0.8313768
	Hessian Percolation	0.9596404	0.8350595	0.8930260
	Hessian Maximal entry filter	0.6088826	0.7225134	0.6608490
Image I_3	Frangi filter	0.9548712	0.6366311	0.7639332
	Sheet filter	0.9091337	0.7658742	0.8313777
	Hessian Percolation	0.9581398	0.7971772	0.8702784
	Hessian Maximal entry filter	0.6803809	0.7617763	0.7187816

Table 1: Precision, Recall and F1-score of methods applied to the input images I_1, I_2, I_3.

Frangi filter with $\sigma = 2.5$ and Hessian Maximal entry filter with $\mathcal{S} = \{1, 3, 5\}$ for a 600^3 CT image I_0, see Figure 5.

The results shown in Figure 5 highlight the distinction between single-scale and multi-scale approaches. While the Frangi filter struggles to classify voxels belonging to thick cracks, the Maximal Hessian Entry filter, with a well-tailored setting for $\mathcal{S}$, can effectively capture various types of cracks with a tolerance for a higher level of noise. Although a multi-scale approach for the Frangi filter is desirable, its computational cost poses a significant obstacle to its practical use for large 3D images.



(a) The input image I_0 (b) Frangi filter with $\sigma = 2.5$ (c) Hessian Maximal entry filter

Figure 5: 2D slices of input images and its outcome produced by Frangi filter and Hessian Maximal entry filter with varying crack widths.

In terms of running time, a comparison was conducted among the Frangi filter, Sheet filter, Hessian percolation and Maximal Hessian Entry filter. Notably, Hessian percolation requires more time due to the need for an initial set H from the binarization phase. For processing 256^3 semi-synthetic images I_1, I_2, I_3 from Figure 3, the Frangi filter and Sheet filter with a single scale take approximately 10-15 seconds, while the Maximal Hessian Entry filter requires less than 3 seconds. With a higher resolution image, such as 600^3, the single-scale approach for the Frangi filter and Sheet filter takes 70-80 seconds, and the multi-scale setting for them requires around 6 minutes. In contrast, the running time for the Maximal Hessian Entry filter is 20 seconds.

Despite the high noise level in the outcomes generated by the Maximal Hessian Entry filter, its advantageous running time and capability to capture the structure of cracks make it a valuable tool for detecting anomaly regions within a reasonable timeframe. Studying the geometry of objects in smaller regions, where cracks predominantly overshadow noise, aids in effectively classifying them. From this point forward, we exclusively utilize the Maximal Hessian Entry filter for further processing.

2.4. Geometry of the binary image $L_{\mathcal{S}}$

In this section, we propose a method to study the geometry of local structure elements, helping to distinguish cracks from noise such as air pores, stones, steel rods, etc. within concrete structures. The key idea is to create a grid of relatively small images out of W, where any existing cracks are expected to exhibit relevant geometry features minimally affected by noise. The proper selection of subimage size is crucial to prevent statistical flattening, a phenomenon that can occur when the size is either too large or too small. Under these conditions, differences in shapes between noise and cracks become reliable contributors for further assessment.

Let $L_{\mathcal{S}} : W \to \{0, 1\}$ represents a binary image. First, we aim to construct a partition of cubes along the axes for the entire concrete structure. Each cube's size is chosen to be neither too small nor too large in comparison to the whole window W, ensuring the capacity to capture objects without statistical flattening, which could lead to information loss. Let $g \in \mathbb{N}$ and $\mathcal{W}_g = \{1, \cdots, g\}^3$, the partition of $L_{\mathcal{S}}$ is defined as

$$W = \bigcup_{q \in \mathcal{W}_g} W(q),$$

where all $W(q)$ share an equal size. This results in a collection of cubes along the axes, denoted as $\{L_q, q \in \mathcal{W}_g\}$, where

$$L_q(p) = \{L_{\mathcal{S}}(p) \in \{0, 1\} : p \in W(q)\}.$$

Second, we compute various geometry statistics from the collection of subimages $\{L_q, q \in \mathcal{W}_g\}$. Intuitively, considering the continuity of cracks and their widths, we focus on subimages with $|W(q)| \in [15^3, 30^3]$. Notably, the crack surface density provides insights into the spatial distribution of cracks, while the volume of foreground in each cube offers a measure of the amount of material present, aiding in distinguishing regions with cracks from those without. Additionally, considering the projection of cracks into planes corresponding to 13 possible 3D grid directions reveals substantial changes in shapes. In contrast, noise such as air pores tends to maintain similar patterns across different directions. Therefore, the standard deviation of areas of projection for 13 different directions emerges as an applicable statistic to differentiate crack regions from noise-dominated areas. Denoting a_q, b_q, and c_q as the surface density, volume of foreground, and standard deviation of areas of projections, respectively, we define $T_q = (a_q, b_q, c_q)$ as a vector containing all statistics obtained from a subimage L_q. The statistics a_q, b_q, and c_q are computed as follows:

$$a_q = \frac{S_q}{V_q}, \quad b_q = V_q, \quad c_q = \mathrm{sd}(S_{q,r}, r \in D)$$

where S_q, V_q, $S_{q,r}$, and D represent the area of the inner surface, volume of foreground, area of projection in direction r, and the parameterized collection of 13 different directions in 3-dimensional space, respectively. This computation is performed using ToolIP software [16]. For more in-depth information on the calculations of V_q and $S_{q,r}$, please refer to [27]. Based on these, we define the following three images:

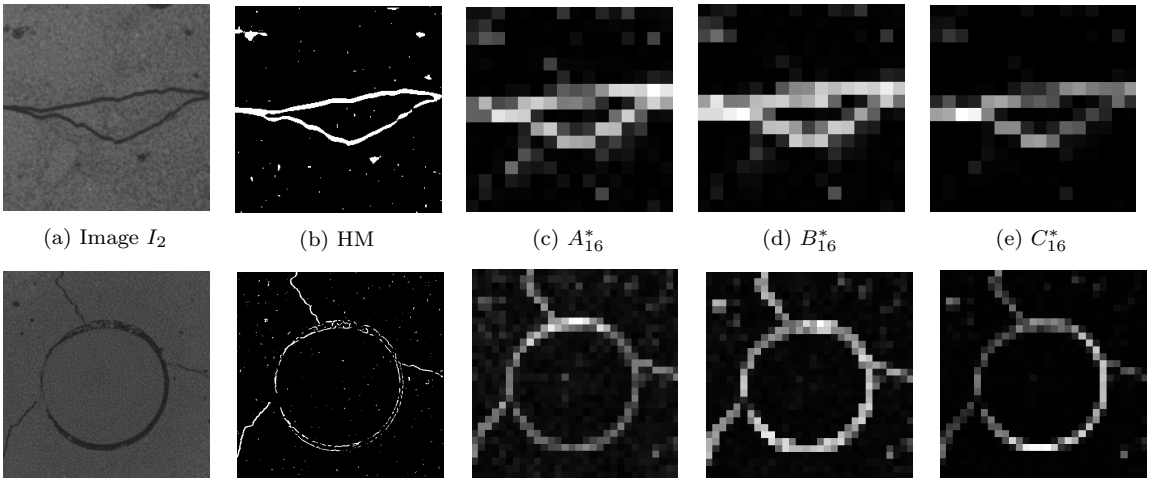
$$A_g = \{a_q, q \in \mathcal{W}_g\}, B_g = \{b_q, q \in \mathcal{W}_g\}, C_g = \{c_q, q \in \mathcal{W}_g\}.$$

It is crucial to observe that the obtained features have different ranges of values. Therefore, it is ideal to standardize each geometry statistic. Let A_g^*, B_g^*, C_g^* be defined as follows:

$$A_g^* = \left\{ \frac{a_q}{sd(A_g)}, q \in \mathcal{W}_g \right\}, \quad B_g^* = \left\{ \frac{b_q}{sd(B_g)}, q \in \mathcal{W}_g \right\}, C_g^* = \left\{ \frac{c_q}{sd(C_g)}, q \in \mathcal{W}_g \right\}.$$

The derived statistics effectively discern disparities between anomaly regions (crack) and homogeneous regions (material). Hence, the images A_g^*, B_g^*, C_g^* ought to replicate a similar crack pattern as the input CT images.

This procedure is applied to both semi-synthetic and real images, compare Figure 6. We choose $g = 16$ for subdividing the 256^3 semi-synthetic image I_2 from Figure 3 and $g = 30$ for the real CT image I_0 from Figure 5.



(a) Image I_2 (b) HM (c) A^*_{16} (d) B^*_{16} (e) C^*_{16}

(f) Image I_0 (g) HM (h) A^*_{30} (i) B^*_{30} (j) C^*_{30}

Figure 6: From left to right: 2D slices of input images, binary image computed from the Maximal Hessian Entry filter and geometric statistic images.

In Figure 6, it is evident that the selected statistics serve as valuable features for identifying regions with cracks. Consequently, one can utilize these features for crack identification, offering an alternative to working with a large-scale input image. Moreover, one can handle all feature images A_g, B_g, C_g as coordinate-wise realizations of a 3-variate random field $\mathcal{X} = \{\xi_k \in \mathbb{R}^3, k \in \mathcal{W}\}$. Within this framework, the classification of anomaly regions essentially involves studying the change-point problem of multivariate random fields.

3. Crack detection as a change-point problem

In the books [11, 12], the change-point problems are discussed under a general parametric setting using a CUSUM statistic test which is widely accepted for anomaly detection, especially in time series; cf. e.g. [31, 4, 23, 9]. For our problem, a change-in-mean CUSUM statistic test will be employed, following the idea proposed in [14, 20].

3.1. Multiple Hypotheses Testing

Let $\mathcal{X} = \{\xi_q, q \in \mathcal{W}\}, \mathcal{W} \subset \mathbb{Z}^n$ denote a stationary, centered, vector-valued random field, $\mathbb{P}\{\xi_q \in \mathbb{R}^d\} = 1$. We introduce the finite parameter space Θ, where for each $\theta \in \Theta$, we define $J_\theta \subset \mathcal{W}$ as the corresponding scanning window within $\mathcal{W}$. The selection of J_θ is crucial, and it should avoid being excessively small or large,

ensuring meaningful analysis. The change-point hypotheses for any $J_\theta, \theta \in \Theta$, are formulated as follows:

$H_0(\theta) : \mathbb{E}\xi_k = \mu \in \mathbb{R}^d$ for all $k \in W$, i.e. there is no change in mean, versus

$H_1(\theta)$: there exists a vector $h \in \mathbb{R}^d, h \neq 0$ such that $\mathbb{E}\xi_k = \mu + h, k \in J_\theta$ and $\mathbb{E}\xi_q = \mu, q \in I_\theta^c$.

To test $H_0(\theta)$ versus $H_1(\theta)$, for $p \geq 1$, employ the change-in-mean CUSUM statistic:

$$T(\theta) = \left\| \frac{1}{|J_\theta|} \sum_{k \in J_\theta} \xi_k - \frac{1}{|J_\theta^c|} \sum_{k \in J_\theta^c} \xi_k \right\|_p, \text{ where } \|\cdot\|_p \text{ is the } L_p\text{-norm in } \mathbb{R}^d.$$

The null hypothesis $H_0(\theta)$ is rejected when the test statistic $T(\theta)$ exceeds a pre-specified critical value denoted as $y(\theta, \alpha)$, with α representing the predetermined significance level. Alternatively, if the p-value $p(\theta)$ associated with the statistic $T(\theta)$ is available, the null hypothesis $H_0(\theta)$ is rejected when $p(\theta) \leq c(\theta, \alpha)$. In order to find the p-value, knowing the null distribution, i.e. the distribution of T_θ) provided H_0 holds is required. However, in the absence of knowledge about the distribution, determining the p-values becomes challenging. To address this issue in practical applications, the empirical null distribution for $T(\theta)|H_0$ can be utilized.

To identify regions containing cracks, a comprehensive approach involves examining all scanning windows, and multiple hypotheses testing is applied. The goal is to test all null hypotheses $H_0(\theta), \theta \in \Theta$, simultaneously. Analogous to single hypothesis testing, where controlling Type-I error and Type-II error at a certain level α are crucial, in the context of multiple testing, there are two types of errors known as family-wise error (FWER) and false discovery rate (FDR). Assuming $|\Theta| = N$ is the total number of hypotheses, let V or S represent the total number of falsely or correctly rejected null hypotheses. The errors FWER and FDR are defined as follows:

$$FWER = \mathbb{P}\{V > 0\} \text{ and } FDR = \mathbb{E}\left[\tfrac{V}{V+S}\right].$$

In other words, FWER signifies the probability of committing one or more false discoveries within a family of tests while FDR is a measure of the expected value of the proportion of false positives among the rejected null hypotheses. The concept of p-values, serving as a quantitative measure of evidence against a null hypothesis, is central to our discussion. In crack detection scenarios with a large number of hypotheses (N), the objective is often to maximize discoveries. Traditional methods controlling FWER can be overly restrictive, leading to conservative tests and reduced power. In such cases, stringent FWER control may miss the detection of true positives. As an alternative, controlling FDR allows for a higher level of false H_0 rejections

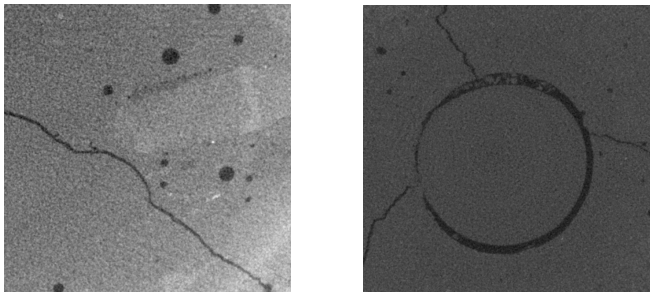
while enhancing test power. This adaptability is crucial in crack detection, where achieving a balance between sensitivity and control reveals meaningful patterns. The well-known Benjamini-Hochberg procedure [7], a step-up approach, effectively controls FDR at a pre-specified level α under the assumption of independence among hypotheses.

Theorem 1 (Benjamini-Hochberg Procedure [7]). *Let $p_{(1)}, p_{(2)}, \ldots, p_{(m)}$ be the ordered p-values from m independent tests. For a given significance level α, reject the null hypotheses corresponding to $p_{(i)}$ for $i = 1, 2, \ldots, k$, where k is the largest index such that $p_{(i)} \leq \frac{i \cdot \alpha}{m}$. This procedure controls the False Discovery Rate (FDR) at level α.*

We will apply this procedure to both semi-synthetic and real concrete images. In the subsequent numerical section, our objective is to illustrate the efficiency of this procedure in localizing anomalies within our dataset. In this paper, we neglect the interdependence of our hypotheses $H_0(\theta)$ vs. $H_1(\theta), \theta \in \Theta$.

4. Numerical results

In this section, our objects are three semi-synthetic images from Figure 3 and two different types of real CT images, provided from the Technical University of Kaiserslautern and Fraunhofer ITWM, see Figure 7.



(a) 500^3 input image I_4 (b) 600^3 input image I_5

Figure 7: 2D slices of real CT concrete images with different sizes.

As mentioned earlier, the theoretical null distribution remains unknown. However, in large-scale simultaneous hypothesis testing, the empirical null distribution can be considered a useful alternative. Additionally, from a practical standpoint, understanding the structure of the material is crucial before applying multiple testing to a

specific type of data. Applying a fixed theoretical threshold for p-values may lead to misclassification when the material's properties fluctuate.

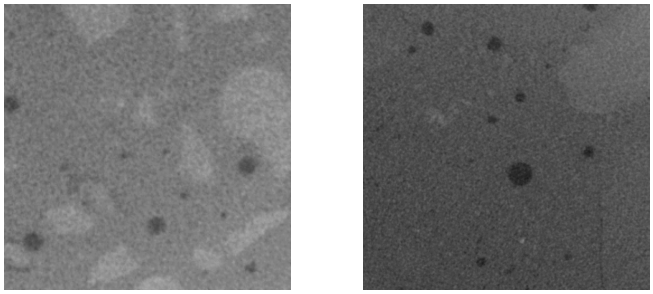


Figure 8: From left to right: 2D slices of 256^3 and 400^3 homogeneous real CT images.

We apply the Maximal Hessian Entry filter with $\mathcal{S} = \{1, 3, 5\}$ to both CT images in 7. The parameter settings for the geometry extraction phase differ. For the 256^3 image, we subdivide it into cubes of size 16^3, resulting in $\mathcal{W} = \{1, 2, \ldots, 16\}^3$. In the case of the 400^3 image, its partition consists of 20^3 cubes, and $\mathcal{W} = \{1, 2, \ldots, 20\}^3$. In the multiple testing step, let $n = d = 3$ and

$$\Theta = \{\theta = (a, b, c, u) : a, b, c \in \mathbb{Z}, J_\theta = [a, a + u] \times [b, b + u] \times [c, c + u] \cap \mathbb{Z}^3 \subset \mathcal{W}\}.$$

For the computation of the CUSUM statistic $T(\theta)$, we use $p = \infty$ for the maximum norm. The scanning window size is denoted by u, and in this instance, we fix $u = 3$. The histograms illustrating the empirical distribution of statistics $T(\theta), \theta \in \Theta$, under H_0 can be found in Figure 9.

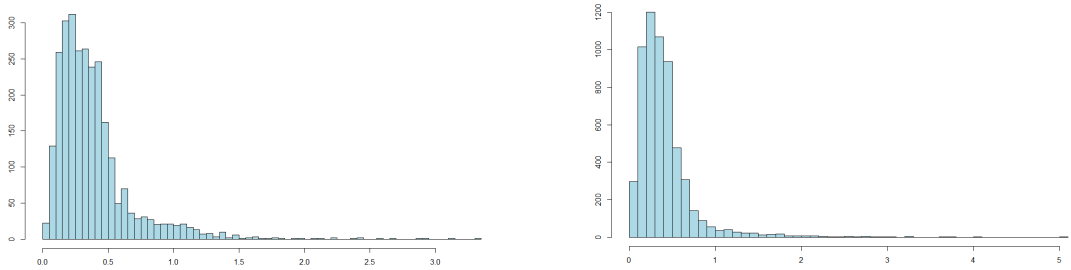


Figure 9: From left to right: Histograms of statistics $T(\theta)$ from 256^3 and 400^3 homogeneous real CT images from Figure 8.

Let $\hat{F}_g$, $g \in \{16, 20\}$ represents the empirical null distribution functions derived from the statistic $T(\theta)$ of 256^3 and 400^3 homogeneous images from Figure 8, respectively. For any computed value of $T(\theta)$, $\theta \in \Theta$, the estimated p-value is defined as

$p(\theta) = 1 - \hat{F}_g(T(\theta))$. Set $R_\alpha(\theta)$ to be 1 if we reject the null hypothesis $H_0(\theta)$ and -1 if we accept it, where α signifies the desired level of False Discovery Rate.

It is crucial to note that due to the overlapping scanning windows, a region within the input image is used in the computations of statistics, thereby the decision that this region has cracks depends on several statistics. An additional step is necessary to precisely identify and visualize the anomaly regions. Recall our partition $W = \bigcup_{q \in \mathcal{W}_g} W(q)$. For each $W(q)$, $\mathcal{I}_q$ represents the collection of scanning windows containing $W(q)$. After completing multiple hypotheses testing, the set $\mathcal{R}_{\alpha,q} = \{R_\alpha(\theta) : J_\theta \in \mathcal{I}_q\}$ is obtained. A region $W(q) \subset W$ is identified to contain a crack if

$$\sum_{x \in \mathcal{R}_{\alpha,q}} x \geq 0.$$

By considering this criterion, all involved scanning windows contribute to making the final decision, leading to higher precision and a reduction in the number of false positives.

4.1. Semi-synthetic images

Consider the 256^3 semi-synthetic images I_1, I_2, I_3, as illustrated in Figure 3, setting parameters $g = 16$, $u = 3$, and $|\Theta| = 2744$, the p-values $p(\theta)$, are estimated using the empirical null distribution $\hat{F}_{16}$, derived from homogeneous data. Multiple testing is performed at two different levels, $\alpha \in \{0.4, 0.5\}$.

To evaluate the classification performance, precision, recall, and F1-score metrics are employed.

	Precision	**Recall**	**F1-score**
$\alpha = 0.4$			
Image I_1	0.1445087	0.9052925	0.2492331
Image I_2	0.7877013	0.7888563	0.7882784
Image I_3	0.8651026	0.3831169	0.5310531
$\alpha = 0.5$			
Image I_1	0.1220044	0.9359331	0.2158689
Image I_2	0.7213115	0.9032258	0.8020833
Image I_3	0.7380952	0.6844156	0.7102426

Table 2: Performance over semi-synthetic data with respect to different levels $\alpha = 0.4, 0.5$ of our multiple testing procedure.

In Table 2, the efficacy of multiple testing is notably influenced by the quality of binary images. The performance, as illustrated in conjunction with Table 1, is

significantly impacted by the low precision in processing the image I_1. This results in an excessively noisy binary image, where noise and pores may be erroneously connected to form a crack. Consequently, the geometric properties become less effective in such scenarios. However, when the noise level is within acceptable limits, the Benjamini-Hochberg procedure demonstrates relatively favorable outcomes in terms of power and precision.

An essential consideration is the choice of the significance level α. Given the unspecified dependency structure among hypotheses, there is an inherent loss of information. The Benjamini-Yekutieli procedure [8] offers a modification to the Benjamini-Hochberg approach, but the threshold for each p-value often tends to make the test conservative. As a result, instead of choosing $\alpha = 0.1$ or 0.2 as recommended, we opt for higher values of 0.4 and 0.5. These adjusted values will be employed in subsequent applications to real-world images.

4.2. Real CT images of concrete

Now, we extend the framework to process real CT input images of concrete samples. Subdividing the images into 20^3 cubes, we set $g = 25$ for 500^3 image I_4, and $g = 30$ for 600^3 image I_5 (cf. Figure 7). The parameter space and significance level α remain the same as in Section 4.1. In these cases, we get $|\Theta| = 12167$ for 500^3 image and $|\Theta| = 21952$ for 500^3 image.

The results in Figure 10 illustrate that our methodology effectively captures the structural characteristics of cracks. The approach demonstrates strong noise elimination from the binary images, benefiting from the detailed geometry analysis of Section 2.4. Notably, selecting a value for α from the range $[0.4, 0.5]$ proves to be particularly effective. However, when dealing with complex crack patterns, the method may face difficulties in accurately detecting anomalies, which could result in a higher false negative rate. This challenge may arise due to the diverse orientations of cracks within a cube, causing geometric statistics to resemble the form of a sphere, commonly associated with air pores or noise. Additionally, the quality of the scale image is a crucial factor. In the case of the 600^3 input image I_5 in Figure 7, variations in lighting conditions may cause certain parts of the crack, expected to exhibit low gray scale values, to behave similarly to homogeneous regions, impacting the detection process.

Additionally, a significant challenge arises from the substantial variation in crack widths. The results heavily depend on the outcomes of the Hessian filter, an edge-detection method that can only partially capture the cracks. This limitation leads to a high rate of false negatives in the subsequent statistics computation and testing. Even when we can capture the crack entirely, determining suitable features for subfields

remains a question.

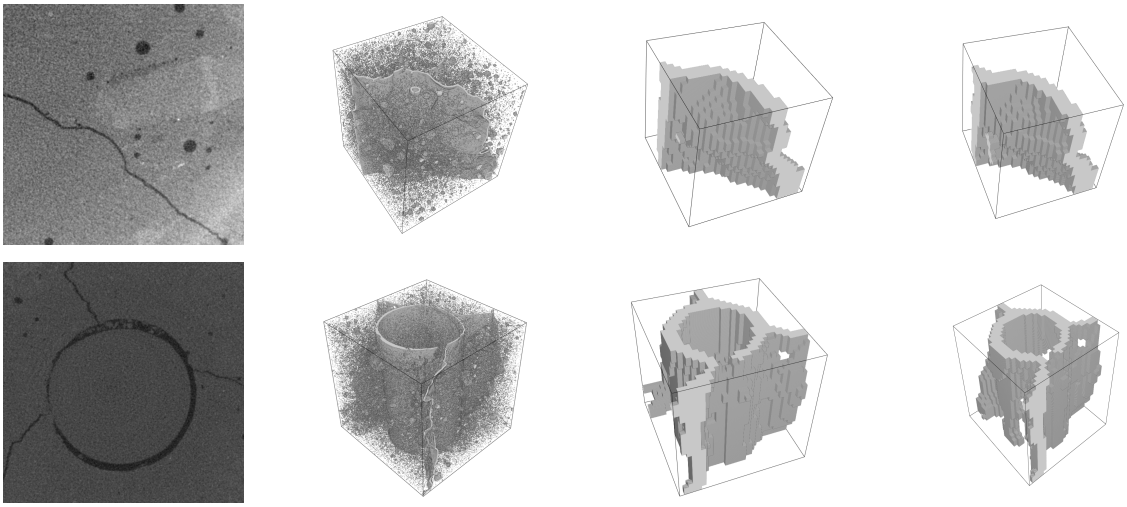


Figure 10: From left to right: Slices of the input image, outcomes produced by Maximal Hessian Entry filter and the predicted anomaly regions with respect to $\alpha = 0.4$ and 0.5, respectively.

4.3. Run time

In this section, we present the runtime analysis of our procedure applied to both semi-synthetic and real concrete images with sizes of 256^3, 500^3, and 600^3. The overall framework involves three main steps: image binarization, geometry study, and performing multiple hypotheses testing.

The experiments were conducted on a desktop PC equipped with an Intel(R) Core(TM) i9-10900K CPU running at 3.70 GHz and 128 GB RAM. For the 256^3 image I_1, I_2, I_3 in Figure 3, the entire process takes approximately 35 seconds, distributed as follows: 3 seconds for applying the Maximal Hessian Entry filter, 30 seconds for geometry study, and less than 2 seconds for the testing procedure. Processing the 500^3 image I_4 from Figure 7 requires around 2.5 minutes, including 14 seconds, 2 minutes, and 5 seconds for each respective step. The complete crack detection analysis for the 600^3 image I_5 from Figure 7 takes approximately 3.5 minutes, with 25 seconds, 3 minutes, and 10 seconds allocated to the individual phases. The most time-consuming task remains evidently in extracting geometric information from each cube L_q. However, since this step is executed sequentially, a substantial reduction in time can be achieved by implementing parallel computing. Analysis of complexity: our method takes $O(|W|)$, where $|W|$ denotes the volume of the image I.

5. Conclusions

This paper introduces a statistical approach for efficient crack localization, addressing the challenges associated with the computational cost in large-scale image processing. Traditional statistical methods for crack detection, especially when employing multi-scale approaches, often incur significant computational expenses. In response to this, our proposed method strikes a balance between computational efficiency and the effectiveness of Hessian-based filters in Section 2.1.

Its key feature is combining a straightforward edge-detection technique with a nuanced analysis of the image geometric structure in subregions, where we control the DFR, i.e. have a specified error probability. This collaborative approach optimizes the performance of the Maximal Hessian Entry filter while managing the computational demands, making it well-suited for large-scale crack detection tasks. The results indicate that our approach achieves the desired goal of crack localization within a reasonable time frame, indicating its practical utility in big data image processing.

It is still a future challenge to test our crack segmentation approach on significantly larger images of $(10000 \times 10000 \times 2000)$ voxels, including those of steel-reinforced concrete. Tackling these challenges requires a robust approach. Once our crack detection algorithm localized potential image regions containing cracks, the subsequent precise crack segmentation methods (including Deep Learning techniques) run on these subregions of a large original image, resulting in drastic runtime reduction and enhancing the effectiveness of crack segmentation.

Funding

This research was funded by the German Federal Ministry of Education and Research (BMBF) [grant number 05M20VUA (DAnoBi)].

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.

References

[1] I. Abdel-Qader, O. Abudayyeh, M. E. Kelly, Analysis of Edge-Detection Techniques for Crack Identification in Bridges, Journal of Computing in Civil Engineering, 17 (4) (2003) 255–263, https://doi.org/10.1061/(ASCE)0887-3801(2003)17:4(255).

[2] P. Alonso-Ruiz, E. Spodarev, Estimation of entropy for Poisson marked point processes, Advances in Applied Probability, 49 (2017) 258–278, http://www.jstor.org/stable/44985422.

[3] P. Alonso Ruiz, E. Spodarev, Entropy-based Inhomogeneity Detection in Fiber Materials, Methodology and Computing in Applied Probability, 20 (2018) 1223–1239, https://doi.org/10.1007/s11009-017-9603-2.

[4] R. Andre-Obrecht, A new statistical approach for the automatic segmentation of continuous speech signals, IEEE Transactions on Acoustics, Speech, and Signal Processing, 36 (1) (1988) 29–40, https://doi.org/10.1109/29.1486.

[5] T. Barisin, C. Jung, F. Müsebeck, C. Redenbach, K. Schladitz, Methods for segmenting cracks in 3D images of concrete: A comparison based on semi-synthetic images, Pattern Recognition, 129 (2022), https://doi.org/10.1016/j.patcog.2022.108747.

[6] T. Barisin, K. Schladitz, C. Redenbach, Riesz Networks: Scale-Invariant Neural Networks in a Single Forward Pass, Journal of Mathematical Imaging and Vision, 66 (2024) 246–270, https://doi.org/10.1007/s10851-024-01171-4.

[7] Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society: Series B (Methodological), 57 (1) (1995) 289–300, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

[8] Y. Benjamini, D. Yekutieli, The Control of the False Discovery Rate in Multiple Testing under Dependency, The Annals of Statistics, 29 (4) (2001) 1165–1188, http://www.jstor.org/stable/2674075.

[9] A. Betken, M. Wendler, Rank-based change-point analysis for long-range dependent time series, Bernoulli, 28 (4) (2022) 2209–2233, https://doi.org/10.3150/21-BEJ1416.

[10] C. E. Bonferroni, Teoria Statistica delle classi e calcolo delle probabilita, Pubblicazioni Del R Istituto Superiore Di Scienze Economiche E Commericiali Di Firenze, 1936.

[11] B. E. Brodsky, B. S. Darkhovsky, Nonparametric Methods in Change Point Problems, Springer Dordrecht, 1993.

[12] B. Brodsky, B. Darkhovsky, Non-parametric statistical diagnosis: problems and methods, Vol. 509, Springer Science & Business Media, 2013.

[13] J. Canny, A Computational Approach to Edge Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (6) (1986) 679–698, https://doi.org/10.1109/TPAMI.1986.4767851.

[14] D. Dresvyanskiy, T. Karaseva, V. Makogin, S. Mitrofanov, C. Redenbach, E. Spodarev, Detecting anomalies in fibre systems using 3-dimensional image data, Statistical Computing, 30 (2020) 817–837, https://doi.org/10.1007/s11222-020-09921-1.

[15] K. Ehrig, J. Goebbels, D. Meinel, O. Paetsch, S. Prohaska, V. Zobel, Comparison of Crack Detection Methods for Analyzing Damage Processes in Concrete with Computed Tomography, International Symposium on Digital Industrial Radiology and Computed Tomography, (2011), https://www.ndt.net/?id=11150.

[16] Fraunhofer ITWM: Department of Image Processing [software], ToolIP, 2024, https://www.itwm.fraunhofer.de/en/departments/bv/products-and-services/toolip.html.

[17] A. F. Frangi, W. J. Niessen, K. L. Vincken, M. A. Viergever, Multiscale vessel enhancement filtering, Lecture Notes in Computer Science, 1496 (2000) 130–137, https://doi.org/10.1007/BFb0056195.

[18] G. L. Golewski, The Phenomenon of Cracking in Cement Concretes and Reinforced Concrete Structures: The Mechanism of Cracks Formation, Causes of Their Initiation, Types and Places of Occurrence, and Methods of Detection - A Review, Buildings, 13 (3) (2023) 765, https://doi.org/10.3390/buildings13030765.

[19] Y. Guo, X. Chen, Z. Wang, Y. Ning, L. Bai, Identification of mixed mode damage types on rock-concrete interface under cyclic loading, International Journal of Fatigue, 166 (2023) 107273, https://doi.org/10.1016/j.ijfatigue.2022.107273.

[20] L. Heinrich, Some Bounds of Cumulants of m-Dependent Random Fields, Mathematische Nachrichten, 149 (1) (1990) 303–317, https://doi.org/10.1002/mana.19901490123.

[21] Y. Hochberg, A sharper Bonferroni procedure for multiple tests of significance, Biometrika, 75 (4) (1988) 800–802, https://doi.org/10.1093/biomet/75.4.800.

[22] S. Holm, A Simple Sequentially Rejective Multiple Test Procedure, Scandinavian Journal of Statistics, 6 (2) (1979) 65–70, http://www.jstor.org/stable/4615733.

[23] R. H. Jones, D. H. Crowell, L. E. Kapuniai, Change Detection Model for Serially Correlated Multivariate Data, Biometrics, 26 (2) (1970) 269–280, https://doi.org/10.2307/2529074.

[24] C. Jung, C. Redenbach, Crack modeling via minimum-weight surfaces in 3D Voronoi diagrams, Journal of Mathematics in Industry, 13 (10) (2023), https://doi.org/10.1186/s13362-023-00138-1.

[25] C. Jung, F. Müsebeck, T. Barisin, K. Schladitz, C. Redenbach, M. Kiesche, M. Pahn, Towards automatic crack segmentation in 3D concrete images, 11th Conference on Industrial Computed Tomography (iCT), 27 (3) (2022), https://doi.org/10.58286/26620.

[26] C. Jung, C. Redenbach, K. Schladitz, VoroCrack3d: An annotated semi-synthetic 3d image data set of cracked concrete, Data in Brief, 54 (2024) 110474, https://doi.org/10.1016/j.dib.2024.110474.

[27] J. Ohser, K. Schladitz, 3D images of materials structures: processing and analysis, in: Measurement of Intrinsic Volumes and Related Quantities, John Wiley & Sons, 2009, pp. 149–194.

[28] S. Ogawa, K. Matsushima, O. Takahashi, Crack Detection Based on Gaussian Mixture Model using Image Filtering, International Symposium on Electrical and Electronics Engineering, (2019) 79–84, https://doi.org/10.1109/ISEE2.2019.8921060.

[29] Z. Šidák, Rectangular Confidence Regions for the Means of Multivariate Normal Distributions, Journal of the American Statistical Association, 62 (318) (1967) 626–633, https://doi.org/10.1080/01621459.1967.10482935.

[30] R. J. Simes, An Improved Bonferroni Procedure for Multiple Tests of Significance, Biometrika, 73 (3) (1986) 751–754, http://www.jstor.org/stable/2336545.

[31] A. Tartakovsky, I. Nikiforov, M. Basseville, Sequential Analysis: Hypothesis Testing and Changepoint Detection, CRC Press, 2014.

[32] C. Westin, Y. Sato, S. Nakajima, N. Shiraga, R. Kikinis, A. Bhalerao, S. Tamura, Tissue Classification Based on 3D Local Intensity Structures for Volume Rendering, IEEE Transactions on Visualization and Computer Graphics, 6 (2) (2000) 160-180, https://doi.org/10.1109/2945.856997.

[33] T. Yamaguchi, S. Hashimoto, Fast crack detection method for large-size concrete surface images using percolation-based image processing, Machine Vision and Applications, 21(5) (2010) 797–809, https://doi.org/10.1007/s00138-009-0189-8.

[34] R. Zhang, Statistical analysis of the non-stationary binomial AR(1) model with change point, Applied Mathematical Modelling, 118 (2023) 152–165, https://doi.org/10.1016/j.apm.2023.01.032.

[35] S. Zhu, X. Xia, Q. Zhang, K. Belloulata, An Image Segmentation Algorithm in Image Processing Based on Threshold Segmentation, Third International IEEE Conference on Signal-Image Technologies and Internet-Based System, (2007) 673–678, https://doi.org/10.1109/SITIS.2007.116.