 doi:10.20944/preprints202107.0154.v1

Article

Numerical simulation based damage identification in concrete

Giao Vu[1], Jithender J. Timothy[1,*], Divya S. Singh[1], Leslie A. Saydak[2,4], Erik H. Saenger[2,3,4] and Günther Meschke[1]

[1] Institute for Structural Mechanics, Ruhr-University Bochum, Universitätsstrasse 150, 44801 Bochum, Germany
[2] Bochum University of Applied Sciences, Am Hochschulcampus 1, 44801 Bochum, Germany
[3] Fraunhofer IEG, Fraunhofer Research Institution for Energy Infrastructures and Geothermal Systems, Am Hochschulcampus 1, 44801 Bochum, Germany
[4] Institute for Geology, Mineralogy and Geophysics, Ruhr-University Bochum, Universitätsstrasse 150, 44801 Bochum, Germany
* Correspondence: timothy.jithenderjaswant@rub.de

Abstract: High costs for the repair of concrete structures can be prevented if damage at an early stage of degradation is detected and precautionary maintenance measures are applied. To this end, we use numerical wave propagation simulations to identify simulated damage in concrete using convolutional neural networks (CNN). Damage in concrete subjected to compression is modeled at the mesoscale using the discrete element method. Ultrasonic wave propagation simulation on the damaged concrete specimens are performed using the rotated staggered finite-difference grid method. The simulated ultrasonic signals are used to train a CNN based classifier capable of classifying three different damage stages (microcrack initiation, microcrack growth and microcrack coalescence leading to macrocracks). The performance of the classifier is improved by refining the dataset via an analysis of the averaged envelope of the signal. The classifier using the refined dataset has an overall accuracy of 90%.

Keywords: microcracking, concrete, feature extraction, damage detection, structural health monitoring, CNN based damage classification.

1. Introduction

Detection of early stage degradation in reinforced concrete structures can considerably reduce costs for repair and retrofitting of the structure. Degradation of concrete subjected to external loads initiates as microcracking around aggregates, pores and defects in the material. With increasing loading, the microcracks start coalescing and eventually form localized macrocracks which promote the ingress of corrosive substances and ultimately may lead to the loss of integrity of the structure. As concrete damage initiates as microcracks, and since these entities are much smaller than the aggregate size, early damage detection is not possible using conventional health monitoring techniques. However, diffuse ultrasonic waves i.e. the multiple-scattered late arriving signals (so-called coda waves [1]) are potentially able to detect weak changes due to multiple sampling of the material. One of the most promising new techniques using coda waves is Coda Wave Interferometry (CWI) [2]. CWI essentially is a method to compare the waves before and after perturbation (due to changes in the microstructure) and has been successfully used for identifying changes in concrete due to a variety of mechanical and environmental loading conditions [3–14]. Recently, [15] used diffusivity of signals obtained from diffuse ultrasounic wave as a feature to quantify microcrack damage induced by alkali-silica reaction and elevated temperature in concrete. With the advances in Machine Learning (ML) techniques and their applications to signal processing (for e.g. speech synthesis [16]), ML methods such as Naive Bayes, Random Forests, Support Vector Machines and Neural networks [17] can be potentially used for identifying damage by training a classifier using a dataset containing labeled coda

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 July 2021 doi:10.20944/preprints202107.0154.v1

waves for various damage states or stress levels. Methods of machine learning have also been applied to the identification of surface cracks with the help of Convolutional Neural Networks using image data ([18–20]).

While there is an agreement that diffuse ultrasonic waves are very sensitive to changes in the micro/mesostructure of concrete subjected to external loadings, reliable identification and classification of the wave measurements with the state of the material, i.e. the level of damage is challenging and far from established science. Recently, numerical simulations are increasingly deployed not only as predictive models but as tools that can be used to virtually test hypotheses and gain new insight. Such a methodology can significantly support experimental observations and provide insights into the plausible limits and potential of coda waves for damage detection (see e.g. [21,22]).

In this paper, we explore a methodology for identifying damage in realistic virtual concrete specimens with the help of ultrasonic waves simulations and discrete element based simulations. Damage in concrete is modeled at the mesoscale (where the coarse aggregates are explicitly resolved), as the principal damage mechanisms (microcracking, crack coalescence and localization) at this scale are driven by the material inhomogeneity [23–25]. The rotated staggered grid is used for performing wave propagation simulations [26]. Given the simulated wave data, we train a convolution neural network to identify the state of damage of the concrete specimen. The accuracy of the classifier is improved by performing feature analysis and data refinement.

2. An overview of the proposed methodology

This section presents a brief summary of the entire workflow illustrated in Figure 1. First a realistic synthetic concrete specimen was generated using the in-house Concrete Mesostructure Generator (CMG) code [27]. Next, a uniaxial compression simulation was performed on the synthetic concrete specimen by means of the Discrete Element Method. The mesostructure snapshots containing information about the spatial distribution and evolution of microcracks are used as input for finite difference simulations of wave propagation with multiple sources and receivers. The time series signals, obtained from the wave propagation simulations, which are the displacement components x, y and z recorded at predefined sensor locations for over 40,000 time steps, are then reformatted (reshaping a vector into an array) as an image before being fed to a CNN classifier. In order to increase the accuracy of the classifier, the averaged envelope was used to reduce the dataset. The performance of the classifier was significantly improved when using the reduced dataset for training the CNN classifier. Note, that in this paper, the word 'coda wave' , 'wave', 'signal' and 'datapoint' refer to the same item and are used interchangeably.

3. Simulation of synthetic concrete mesostructures subjected to uniaxial compressive loads

In this section, synthetic concrete mesostructures subjected to various levels of compressive loads are generated. The morphology of plain concrete is characterized by a distribution of coarse and fine aggregates (up to a volume fraction of 60 - 70%) that are embedded in a cementitious matrix (host) material. When concrete specimens are subjected to external loads, these aggregates induce highly disordered and complex stress and displacement fields. Under increasing loads, microcracks first initiate in the vicinity of the aggregates. These microcracks later coalesce to form localized macrocracks [28]. Hence, in order simulate damage processes, especially the pre-peak diffuse damage due microcracking, it is essential that the heterogeneity of concrete at the mesoscale and the distribution of the aggregates are taken into account. To this end, the open-source Concrete Mesostructure Generator (CMG, ([29])) developed by the authors was used to generate an ultra-realistic concrete specimen in which the coarse aggregates are explicitly resolved. The physical dimensions of the synthetic concrete specimen is $10 \times 10 \times 40$ cm. Voxel resolution of the specimen is $200 \times 200 \times 800 = 32$ milion voxels. Each voxel is of size 0.5 mm. The total volume fraction of coarse aggregates in this specimen is 31.11 % with $D_{min} = 15$ mm and $D_{max} = 30$ mm.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 July 2021 doi:10.20944/preprints202107.0154.v1

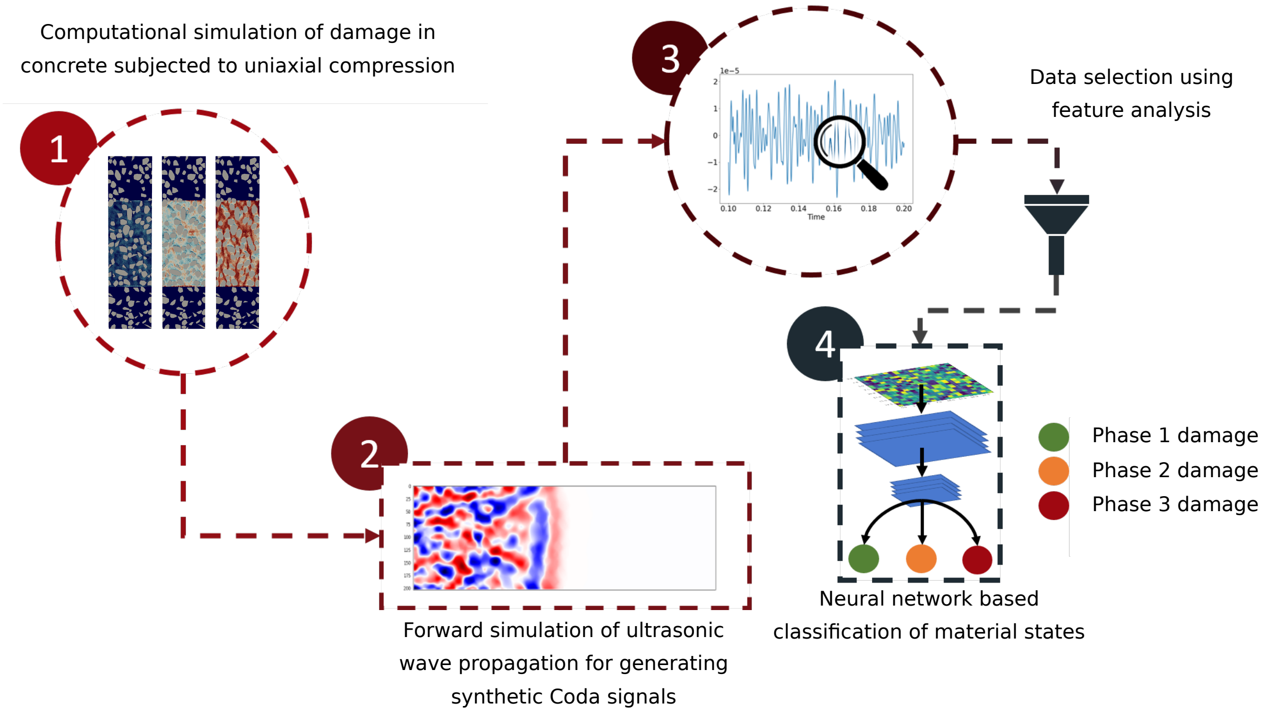

Figure 1. Workflow of the simulation based procedure to detect damage in concrete specimens subjected to uniaxial compressive loadings.

In order to simulate the load induced morphological and material changes to the concrete specimens, we use the Discrete Element Method (DEM). According to DEM, a granular material (such as concrete, rock) is described by a dense packing of discrete elements that are linked together by cohesive frictional forces. Within the medium, the induced forces are transmitted via a contact network between the discrete elements and the dynamics of these discrete elements is governed by Newton s second law [30]. The contact network is first established and updated by identifying the elements and their nearest neighbor interactions. The interaction forces are evaluated based on the relative displacements in the current configuration. Next, the resultant interaction forces are used in conjunction with the applied external forces as input for the equations of motion in the time integration step to solve for the new position of all discrete elements. The concrete constitutive model proposed in [31] is adopted in this work. The interaction of the particles under tension in normal direction is governed by a linear damage softening law and the tangential behavior of contact is based on the modified Mohr-Coulomb frictional law. Details of contact description can be found in [32].

Table 1: Summary of model parameters required for the calibration of the DEM model.

	Elastic parameters	Mortar	Aggregate	
K_n	normal modulus	8	16	GPa
K_τ	tangential modulus	1	2	GPa
	Damage law in tension			
ε_0	limit elastic strain	1×10^{-4}	1×10^{-4}	
$\frac{\varepsilon_f}{\varepsilon_0}$	relative ductility	50	50	
	Elasto-plasticity in shear			
c_0	initial cohesion	1	2	MPa
$\tan\phi$	frictional angle	0.57	0.57	

 doi:10.20944/preprints202107.0154.v1

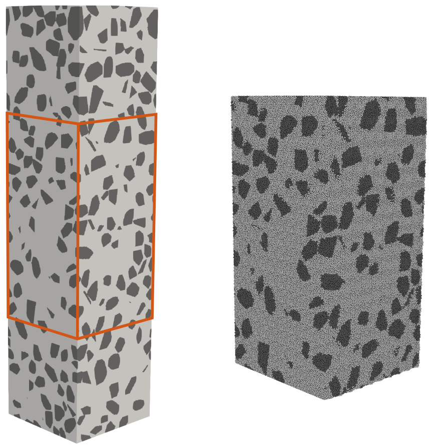

Figure 2. Voxelized concrete sample (left) and the corresponding DEM sample (right) ($d_{min} = 15$ mm, $d_{max} = 30$ mm) considered in the uniaxial compression simulation.

The DEM model is resolved using 2.4 million discrete elements with radius 0.5 mm. Next, the voxelized mesostructure was used to assign the material properties to all discrete elements. Here, only a region in the center of the voxelized mesostructure is considered (z = 10 to 30 cm) and denoted as "affected region", as highlighted by the orange frame in Figure 2. The contact parameters considered in the uniaxial compression simulation are listed in table 1.

From the DEM simulation, a total of 12 mesostructure snapshots are selected and grouped into 3 labels (see Figure 3) "Phase 1 damage" (green colored markers),"Phase 2 damage" (yellow colored markers) and "Phase 3 damage" (red colored markers). Figure 3 (top right) shows the normalized compressive stress (σ_c) vs. the strain obtained from the DEM simulation. Let f_c denote the uniaxial compressive strength of the specimen. In the regime denoted as "Phase 1 damage" ($\sigma_c / f_c = 0 - 0.544$), microcracks start initiating in the specimen, however, the deformation predominantly is elastic at this early stage (see Figure 3). The regime "Phase 2 damage" ($\sigma_c / f_c = 0.691 - 0.937$) is characterized by the growth of diffusely distributed microcracks, also called as "failure precursors". "Phase 3 damage" is related to the post-peak regime associated with multiple macrocracks opening parallel to the loading direction due to the coalescence of the microcracks.

Table 2: List of mesostructure snapshot ids and the corresponding labels and strain levels.

Label	Snapshot id	Strain level [%]
Phase 1 damage	0	0.
	1	0.031
	2	0.062
	3	0.093
Phase 2 damage	4	0.124
	5	0.155
	6	0.186
	7	0.217
Phase 3 damage	8	0.233
	9	0.248
	10	0.264
	11	0.279

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 July 2021 doi:10.20944/preprints202107.0154.v1

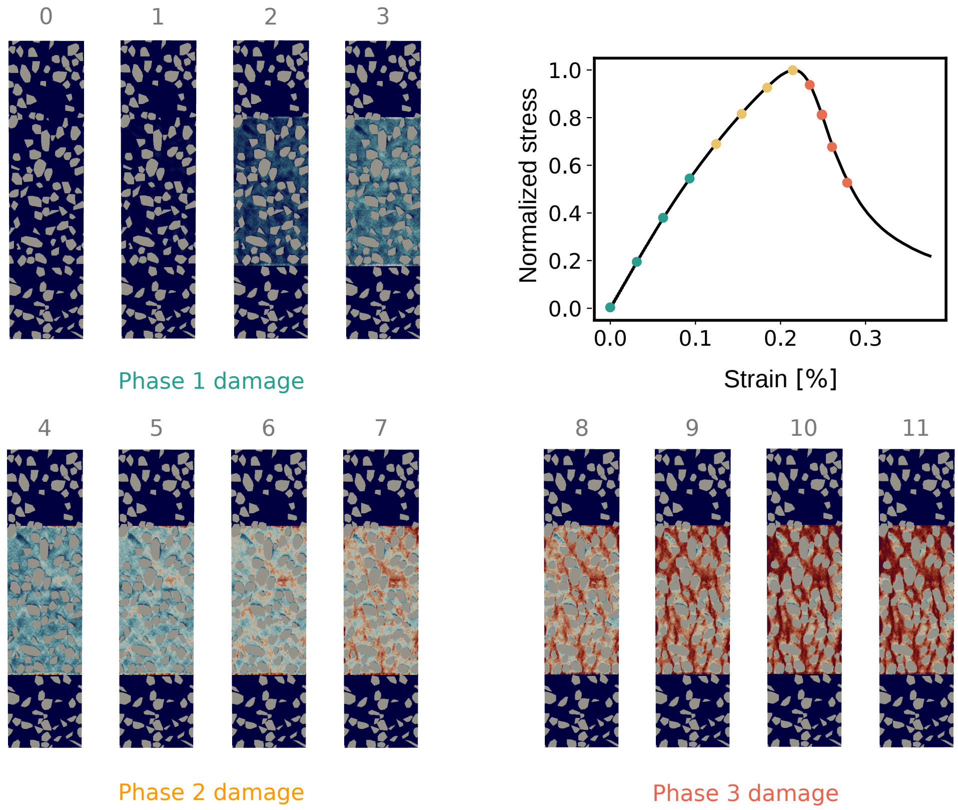

Figure 3. Damage evolution of the concrete sample subjected to uniaxial compression. Legend: grey color - aggregate, blue color - undamaged mortar, red color - damaged mortar. Top right: normalized stress-strain curve obtained from the DEM simulation. The corresponding load levels at which the microstructure snapshots are extracted are highlighted by colored dots (see Table 2 for the strain levels of each damage snapshot).

4. Synthetic coda wave generation

Numerical wave propagation simulations are performed on each of the 12 mesostructures. A rotated staggered-grid (RSG) finite-difference scheme [26] is used to propagate the seismic wavefield in the forward simulations. The RSG uses rotated finite-difference operators, leading to a distribution of modeling parameters in an elementary cell where all components of one physical property are located only at one single position. This can be advantageous for modeling wave propagation in anisotropic media or complex media, including high-contrast discontinuities, because no averaging of elastic moduli is needed [33]. Concrete is a strongly heterogeneous and densely packed composite material. Due to the high density of scattering constituents and inclusions, ultrasonic wave propagation in this material consists of a complex mixture of multiple scattering, mode conversion and diffusive energy transport. With previous studies in 2D and 3D [34,35] it was shown that the RSG-technique is well-suited for such applications. A finite-difference operator of second order is used in time as well as in space. The surfaces of the concrete model are implemented as a free surface using two layers of vacuum. These vacuum layers mimic the impedance of the concrete specimen-atmosphere boundary. A typical simulation with a model size of $40 \times 10 \times 10$ cm^3 consists of 32 million grid points (spatial increment 5×10^{-4} m). A simulation with $40,000$ time steps takes about 3 hours on three nodes on a mid-size cluster computer.

4.1. Simulation set up and time series dataset acquisition procedure

For the mortar matrix we use a P-wave velocity of $v_p = 3950$ m/s, a S-wave velocity of 2250 m/s and a density of 2050 kg/m^3 and for the aggregates a P-wave velocity of $v_p = 6230$ m/s, a S-wave velocity of 3330 m/s and a density of 2950 kg/m^3 is assumed.

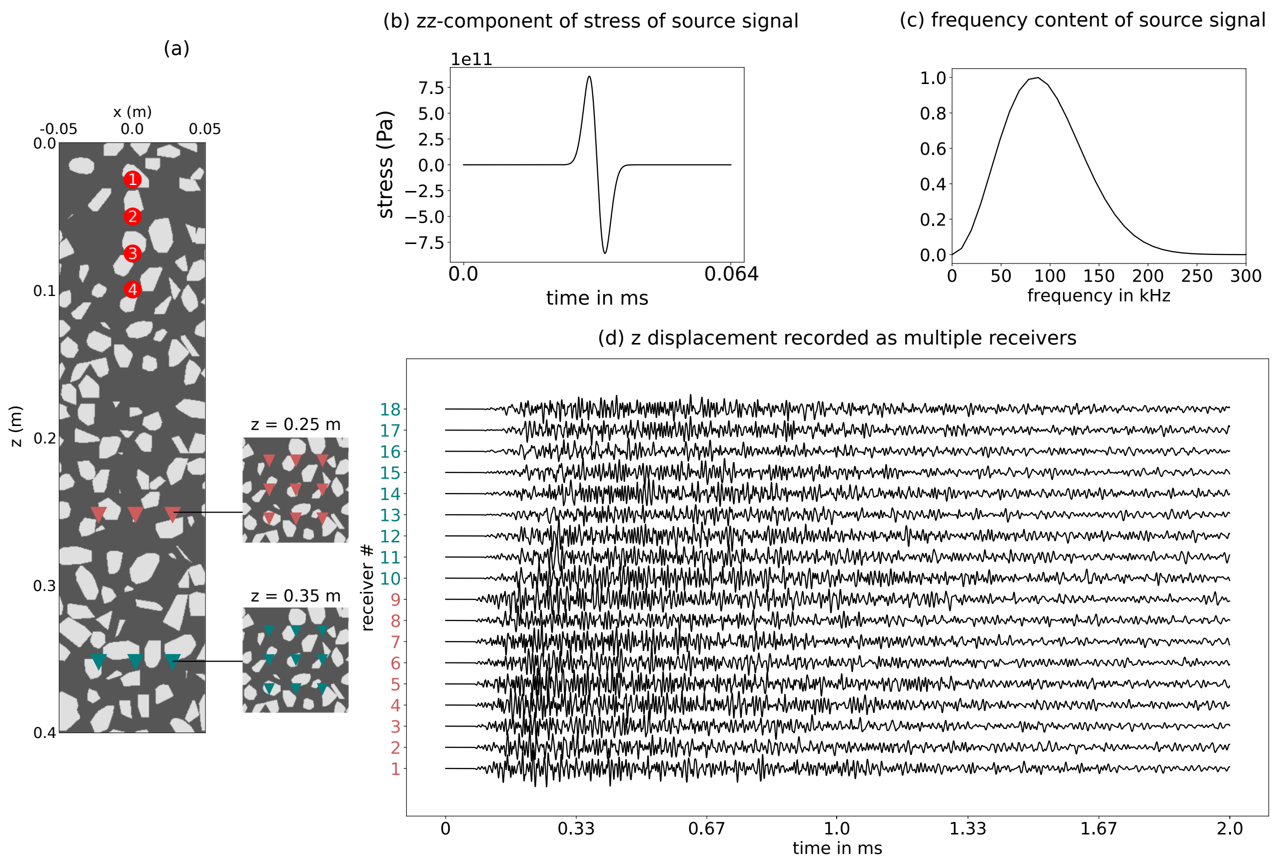

Figure 4. a) Sender-Receiver configuration specified for the wave propagation simulation. b) zz-component of stress of source signal and c) frenquency content of the source signal. d) signals recorded at 18 sensor locations specified in a).

To mimic the attenuation characteristics typically observed in ultrasonic laboratory measurements of concrete specimens, a homogeneous intrinsic attenuation has been applied in the Finite Difference scheme according to [36], with the parameters $Y_1^{11} = 916$ MPa, $Y_1^{44} = 262$ MPa MPa for the aggregates, and $Y_1^{11} = 256$ MPa, $Y_1^{44} = 83$ MPa for mortar matrix. Also, a frequency-dependent attenuation with a single Maxwell body with a minimal $Q = 250$ at $f_{fund} = 150$ kHz is simulated.

For the numerical samples experiencing material damage, the elastic wave velocities are reduced based on the damage parameter d. For all simulations, a wavelet with a central frequency of $fc = 85kHz$ is used, as this central frequency lies within the proposed frequency range for concrete with an ideal signal-to-noise ratio (ISO 1920-7:2004). Sources are implemented as body force sources, with components xx and zz being nonzero (Fig. 4 b,c). The time increment is set to 5×10^{-8} s to ensure numerical stability. In all simulations, multiple sources and receivers were implemented as follows:. Four source locations were considered in the analysis at positions $z = 2.5$ cm, 5 cm, 7.5 cm and $z = 10$ cm.

Fig. 4a shows the setup for the wave propagation simulation. The red circles denote point sources. At positions $z = 25$, and 35 cm, two sets of receivers (red and green triangles) are distributed evenly on a 3-by-3 grid on the xy-plane with a spacing of 2.5 cm. The receivers are denoted with ID:identifiers from 1 to 18. The receivers 1 to 9 (ID = 1–9) lie inside the damaged zone and the receivers with ID = 10–18 lie outside of the damaged zone. The data recorded are the displacement components x, y, z for over 40,000 time steps (i.e. for a total duration of 2 ms). Fig. 4 shows an example of temporal displacements z recorded at 18 locations (also called as time series) corresponding to the first mesostructure shown in Figure 3. In summary, from a single source and 18 receivers, we obtain a total of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 July 2021 doi:10.20944/preprints202107.0154.v1

18 ×3 displacement components ×12 damage cases = 648 time series signals. Considering 4 source locations, a dataset of 2592 simulated time series signals is generated.

5. Damage identification using Supervised Machine learning

Having generated the dataset containing 2592 labeled data points (coda signals), the objective now is to use supervised learning to identify the state of a concrete specimen given a coda signal. The possible states of the concrete specimen correspond to three labels associated with each data point (coda signal) i.e. 1) Phase 1 damage, 2) Phase 2 damage, 3) Phase 3 damage. We use a feed forward Convolutional Neural Network (CNN) based classifier. CNNs are a special type of Artificial Neural Networks (ANN) which are designed specifically to handle inputs with large dimensions. Traditionally, CNNs were designed to classify conventional visual images, in which the pixel data of an image is fed into the network in terms of an array. In our case, the time series data was reshaped into a 2D matrix by first slicing the series-data into multiple parts and then arranging these parts into an array consecutively one row after another.

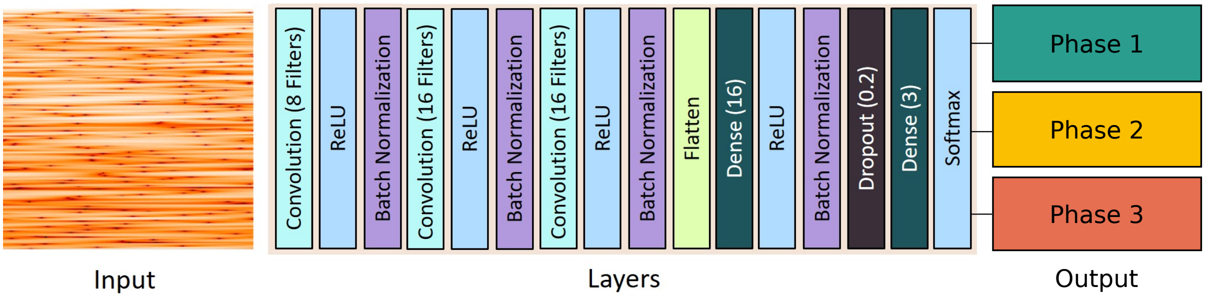

Figure 5. CNN architecture. The input is presented in the form of an image. A total of 2592 input images are used for training the network. There are three convolution layers followed by flattening, dense, dropout and dense. Softmax has been added in the end to convert the output in terms of probability.

A single complete time series signal in our case consists of 40,000 displacement values (corresponding to 40,000 time steps). Reshaping this one dimensional array into a two dimensional array produces an array of size 200 × 200. The output is in the form of a one-hot encoding vector corresponding to each damage phase (i.e. the first label is [1,0,0], the next one is [0,1,0] and the final one is [0,0,1]). The loss function for back propagation is based on Categorical Cross Entropy function. The learning rate was set as 0.001 and the network was trained for 200 epochs with a batch size of 32.

To take into account the influence of signals at the inital state (corresponding to snapshot id 0), all signals of the respective source location is scaled by dividing the time series to the maximum amplitude of the reference signal. Thus, the reference time series signals (displacement components x, y, z at the zero-stress state) have the maximum values between −0.5 and 0.5, and signals at the remaining steps are scaled appropriately.

The dataset (2592 signals) was randomly shuffled, and 3/4 (1944 signals) of the shuffled dataset was used for training and 1/4 (648 signals) of the dataset was used for validation. The NADAM algorithm was used as the optimizer and ReLU was used as the activation function across the network. This optimizer and the activation function was chosen because it provided the fastest convergence with highest accuracy among all optimizers available in the Keras library for this dataset (see Figures A1 and A2, in Appendix A). Figure 8 shows the training and validation accuracy. The validation accuracy is around 77%. The confusion matrix of the trained classifier is shown in Figure 8.

The performance of the classifier using this architecture and the dataset is as follows:

1. Phase 1 damage: 85 percent of the signals corresponding to this phase were correctly identified. 15 percent of the signals were wrongly identified as belonging to phase 2 damage.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 July 2021 doi:10.20944/preprints202107.0154.v1

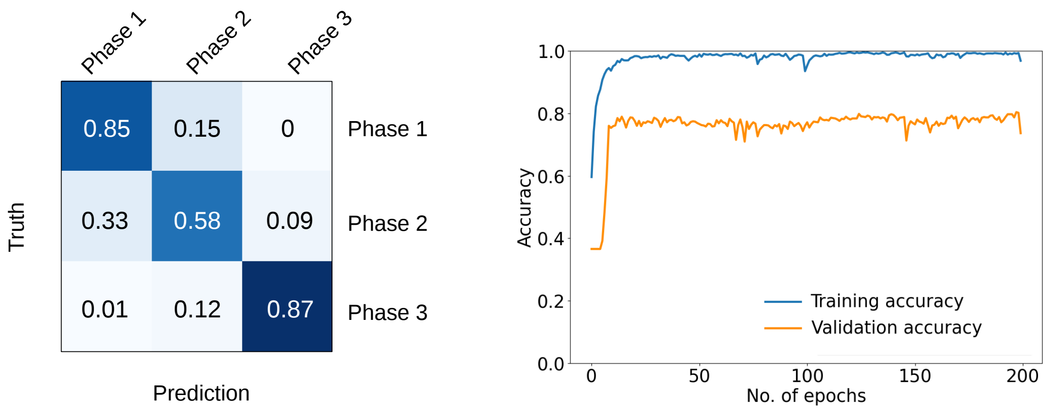

Figure 6. Left: Confusion matrix visualizing the performance of the classifier as the percentage of correct and incorrect classification of the damage phase (three phases of damage (Phase 1: Elastic deformation and microcrack initiation, Phase 2: pre-peak microcracking and Phase 3: Microcrack coalescence and post-peak crack localization)). Right: Overall training and validation accuracy of the classifier for 200 epochs.

2. Phase 2 damage: Only 58 percent of the signals corresponding to phase 2 damage i.e. pre-peak microcracking were correctly identified. Indeed 33 % of the signals belonging to the class of phase 2 damage were incorrectly classified as phase 1 and 9 % were incorrectly identified as phase 3.
3. Phase 3 damage: 87 percent of the signals corresponding to this state were correctly identified. 12 percent were misclassified as phase 2 and 1 percent was wrongly classified as phase 1.

Table 3: Detailed architecture of Convolutional neural network for damage classification

Layer (type)	Output Shape	Filter Size	Stride	No. of Parameters
Conv 2D	[98,98,8]	[5,5]	[2,2]	208
Batch Normalization	[98,98,8]	-	-	32
Conv 2D	[47,47,16]	[5,5]	[2,2]	3216
Batch Normalization	[47,47,16]	-	-	64
Conv 2D	[45,45,16]	[3,3]	[1,1]	2320
Batch Normalization	[45,45,16]	-	-	64
Flatten	[32400]	-	-	0
Dense	[16]	-	-	518416
Dropout	[16]	-	-	0
Dense	[16]	-	-	272
Batch Normalization	[16]	-	-	64
Dense	[3]	-	-	51
Total number of parameters : 524707				

6. Improvement of accuracy using data refinement

In order to improve the classification accuracy especially with regards to the correct identification of damage phase 2 i.e. pre-peak microcracking of concrete, we analyze in this section an additional feature corresponding to the averaged envelope of the coda signal. Using this additional feature we generate a refined dataset. The averaged envelope of all the coda signals from the dataset (2592 signals) was computed by first finding the peaks in the signal over the complete signals and averaging these values. These averaged values were first segregated into different groups corresponding to each source location. These values were then normalized between 0 and 1 corresponding to each source location (the highest value in the group being set to 1 and lowest value as 0 and all the values in between scaled appropriately) before being plotted as a scatter plot.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 July 2021 doi:10.20944/preprints202107.0154.v1

Figure 7 shows 4 different scatter plots of the normalized amplitude across the three damage phases. Each plot corresponds to a particular source location. The data in each plot is segregated according to the location of the receivers. A red background is applied to all data recorded by the sensors 1-9 (lie inside the damaged zone) and a green background is applied to all data recorded by the sensors 10-18 (see Figure 5 for the sensor positions). The average value of this feature for a particular source is plotted as a function of increasing strain levels as a dashed line (horizontal) in Figure 7. Furthermore, two vertical dash lines are placed in the plot to separate the three phases of damage. Two main conclusions can be drawn from the plot. Firstly, this feature corresponding to the sensors inside the damaged zone (sensors 1–9) shows a high variance in comparison to the sensors present outside the damaged zone (sensors 10–18) for all source locations. Secondly, the excitation source (source at 10 cm, bottom right of Figure 7), located inside the damage zone, shows a unique pattern as compared to the other three cases, when the source was located outside the damaged zone. These observations are explained in more detail in the following section.

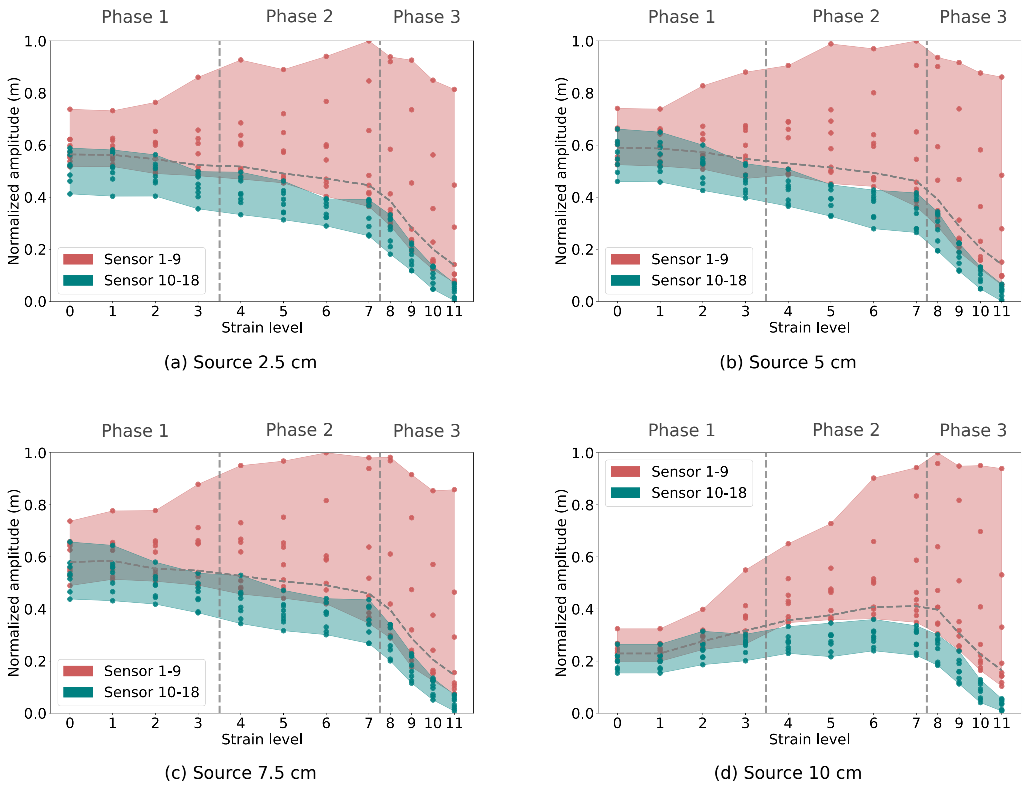

Figure 7. Scatter plots of the average of the averaged envelope of the Coda signals as a function of increasing strain levels (from level 0 to 11) obtained for four locations of the sender (see Table 2 for the exact values of the strains corresponding to the strain levels). In each plot, three phases of damage (Phase 1: Elastic deformation and microcrack initiation, Phase 2: pre-peak microcracking and Phase 3: Microcrack coalescence and post-peak crack localization) of the concrete specimen are indicated.

When the excitation point was located at 2.5 cm (Figure 7, top-left) from the origin along the Z axis, the average of the features for this particular source location reduced with increasing damage. A similar trend is seen for the sources located at 5 cm and 7.5 cm, respectively. This trend can be explained as follows: when the strain level increases, the number of small microcracks present in the structure also increases. A higher number of microcracks present in the path of the source and sensor will reduce the energy present in the wave reaching the sensor because some part of the wave will be reflected back and only a portion will proceed forward at each microcrack. Thus, there is a general decrement in the average envelope value of the Coda signals with increasing strain level.

However, the trend observed for source location 10 cm (Figure 7, bottom-right), which is located inside the damaged zone, is quite different. Here, the line of the average amplitude increases till the peak stress (end of Phase 2) and then decreases afterwards in the post-peak regime. Obviously, the presence of the source location inside the damaged zone affects the trend of the average line. In particular, in the plot from source 10 cm, we observed that, the overall increase of the average amplitude from the red area before the crack localization phase is more pronounced. Interestingly, the averaged amplitude of signals from the green area also increases slightly prior to the crack localization. This might be related to the crack orientation parallel to the sender-receiver orientation, which guides the waves toward the receivers. As long as the damage level is weak, this effect seems to overcompensate the effect of stiffness reduction which is the dominant effect in the other three plots. As soon as crack localization starts (phase 3), the strong stiffness degradation finally governs the decrease of the amplitudes. From our interpretation, this could be a feature, in addition to differentiating diffuse microcracking and localized fracture, could provide information of the damage anisotropy (microcrack orientation). Nevertheless, the potential use of this observation for damage identification still has to be validated according to the analysis of measured coda signals in the laboratory. This will be performed in the future.

When comparing the effect of sensor location, it can be observed in each graph in (Figure 7) that the variance is evidently higher for the group of sensors present inside the damaged zone. It is quite clear that the location of the sensors has a significant effect on the recorded displacement data. The envelope plots provide a simple method to differentiate between "good" and "bad" data. Good data refers to data with a small variance relative to the average line, while bad data refers to data with a large variance. Therefore, we have filtered out the data corresponding to the sensor locations 1 to 9 and the source 10 cm located inside the damage zone. Thus, the training dataset reduces to 972 time series signals. Figure 8 left shows the confusion matrix and Figure 8 right shows the accuracy of the training and validation obtained from the full and the reduced dataset, using the same CNN architecture. We observed, that the accuracy of the validation with the reduced dataset is approximately 90%, which is a significant improvement as compared to taking the full data set, where an accuracy of approximately 77 % was obtained and unsatisfactory prediction of phase 2 damage (58% accuracy).

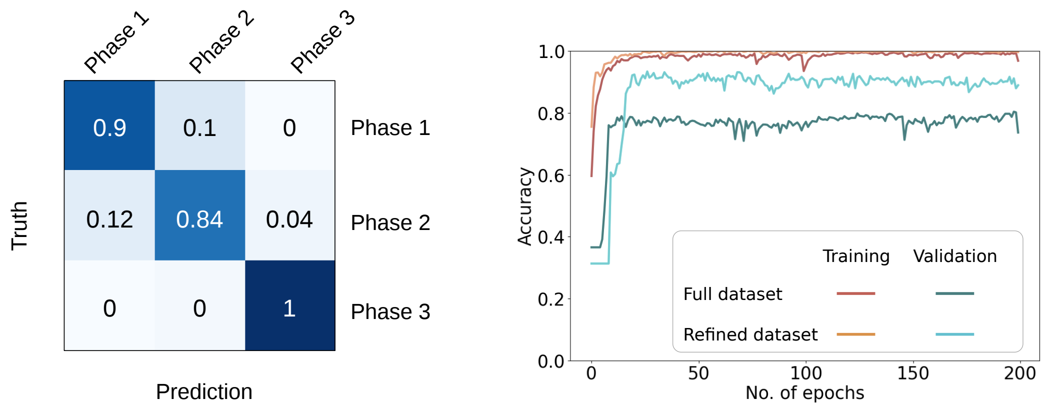

Figure 8. Left: Confusion matrix visualizing the performance of the classifier using the refined dataset. The matrix shows the percentage of correct and incorrect classification of the damage phase (three phases of damage (Phase 1: Elastic deformation and microcrack initiation, Phase 2: pre-peak microcracking and Phase 3: Microcrack coalescence and post-peak crack localization)). Right: Overall training and validation accuracy of the classifier for 200 epochs.

7. Conclusion

A computational methodology for detection of damage in concrete under compression loading using synthetic Coda waves has been proposed in this paper. The microcrack

 doi:10.20944/preprints202107.0154.v1

growth evolving in realistic virtual concrete specimens, characterized by a realistic aggregate size distribution was simulated using the discrete element method (DEM). The damage evolution in the pre-peak regime as well as the formation of localized macrocracks in the post-peak regime have been recorded as different stages and used for computational wave propagation simulations. The simulated ultrasonic signals were used to train a convolutional neural network (CNN) to predict the corresponding damage class (Phase 1: Elastic deformation and microcrack initiation, Phase 2: Pre-peak microcracking and Phase 3: Microcrack coalescence and post-peak crack localization). It was observed that the performance of the CNN using the full dataset was unsatisfactory. Therefore, a feature based on the envelope of the coda signals was used to refine the dataset. The trendline of the average envelope plotted over the strain level shows a remarkable difference when evaluated for the sender located inside the damage zone as compared to the source located outside of the damage zone of the virtual specimen. This feature can potentially be used to derive information on the microcrack orientation. It was also observed, that a much larger scatter of the average of the envelopes of the Coda signals was obtained for sensors located inside the damaged area as compared to the sensors located outside of this zone. When filtering the signals from receivers located inside the damaged zone, i.e. using a reduced data set, the accuracy of the CNN with the same architecture was considerably improved.

 doi:10.20944/preprints202107.0154.v1

229 ## Appendix A Analysis of activation function and optimizer

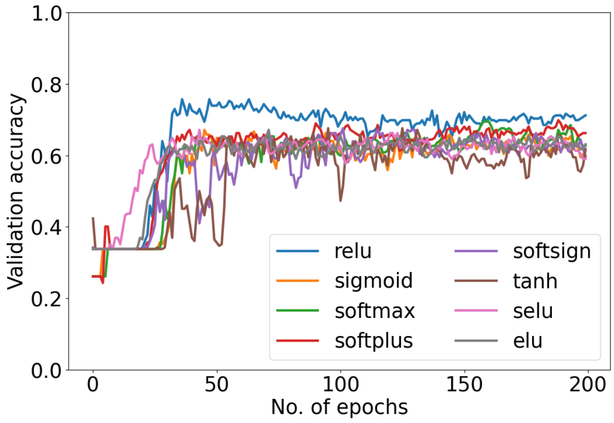

Figure A1. Effect of different activation functions on the validation loss of the CNN network

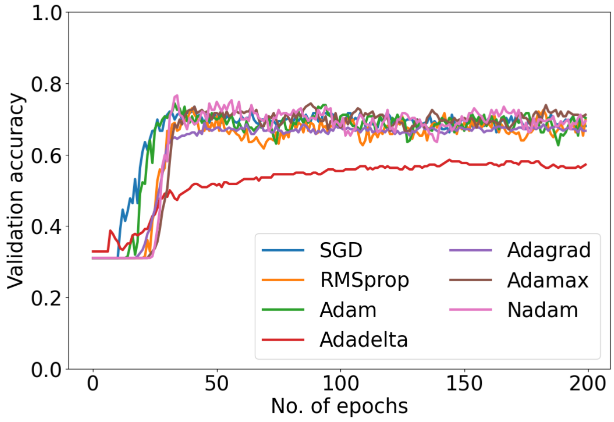

Figure A2. Effect of different optimizers on the validation accuracy of the CNN network

Author Contributions: Conceptualization, G.V., L.S, J.J.T., E.H.S, and G.M. ; methodology, all; software, G.V., L.S, D.S; formal analysis, G.V, D.S, L.S.; investigation, G.V, D.S, L.S, J.J.T; resources,J.J.T, E.H-S and G.M ; data curation, G.V and L.S ; writing—original draft preparation, D.S, G.V, J.J.T; writing—review and editing, All; visualization, G.V D.S and J.J.T; supervision, J.J.T, E.H.S and G.M; project administration, J.J.T, E.H.S and G.M; funding acquisition, J.J.T, E.H.S and G.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the German Research Foundation (DFG) in the framework of Subprojects RUB1 and BU of the Research Unit FOR 2825 (project number 398216472).

Conflicts of Interest: The authors declare no conflict of interest.

The following abbreviations are used in this manuscript:

DEM	Discrete Element Method
RSG	Rotated Stagger Grid
CNN	Convolutional Neural Network
CWI	Coda Wave Interferometry
ML	Machine Learning
CMG	Concrete Mesostructure Generator
ANN	Artificial Neural Network
ReLU	Rectified Linear Unit
SGD	Stochastic Gradient Descent

References

1. Herraiz, M.; Espinosa, A. Coda waves: a review. *Pure and applied geophysics* **1987**, *125*, 499–577.
2. Snieder, R. The theory of coda wave interferometry. *Pure and Applied geophysics* **2006**, *163*, 455–473.
3. Planès, T.; Larose, E. A review of ultrasonic Coda Wave Interferometry in concrete. *Cement and Concrete Research* **2013**, *53*, 248–255.
4. Stähler, S.C.; Sens-Schönfelder, C.; Niederleithinger, E. Monitoring stress changes in a concrete bridge with coda wave interferometry. *The Journal of the Acoustical Society of America* **2011**, *129*, 1945–1952.
5. Zhang, Y.; Abraham, O.; Tournat, V.; Le Duff, A.; Lascoup, B.; Loukili, A.; Grondin, F.; Durand, O. Validation of a thermal bias control technique for Coda Wave Interferometry (CWI). *Ultrasonics* **2013**, *53*, 658–664.
6. Wang, X.; Niederleithinger, E. Coda Wave Interferometry used to detect loads and cracks in a concrete structure under field conditions. Proceedings of the 9th European Workshop on Structural Health Monitoring Series, Manchester, UK, 2018, pp. 10–13.
7. Niederleithinger, E.; Wang, X.; Herbrand, M.; Müller, M. Processing ultrasonic data by coda wave interferometry to monitor load tests of concrete beams. *Sensors* **2018**, *18*, 1971.
8. Legland, J.B.; Zhang, Y.; Abraham, O.; Durand, O.; Tournat, V. Evaluation of crack status in a meter-size concrete structure using the ultrasonic nonlinear coda wave interferometry. *The Journal of the Acoustical Society of America* **2017**, *142*, 2233–2241.
9. Schurr, D.P.; Kim, J.Y.; Sabra, K.G.; Jacobs, L.J. Damage detection in concrete using coda wave interferometry. *Ndt & E International* **2011**, *44*, 728–735.
10. Deraemaeker, A.; Dumoulin, C. Embedding ultrasonic transducers in concrete: A lifelong monitoring technology. *Construction and Building Materials* **2019**, *194*, 42–50.
11. Hafiz, A.; Schumacher, T. Monitoring of stresses in concrete using ultrasonic coda wave comparison technique. *Journal of Nondestructive Evaluation* **2018**, *37*, 1–13.
12. Shokouhi, P. Stress-and damage-induced changes in coda wave velocities in concrete. AIP Conference Proceedings. American Institute of Physics, 2013, Vol. 1511, pp. 382–389.
13. Hilloulin, B.; Legland, J.B.; Lys, E.; Abraham, O.; Loukili, A.; Grondin, F.; Durand, O.; Tournat, V. Monitoring of autogenous crack healing in cementitious materials by the nonlinear modulation of ultrasonic coda waves, 3D microscopy and X-ray microtomography. *Construction and Building Materials* **2016**, *123*, 143–152.
14. Clauß, F.; Epple, N.; Ahrens, M.A.; Niederleithinger, E.; Mark, P. Comparison of Experimentally Determined Two-Dimensional Strain Fields and Mapped Ultrasonic Data Processed by Coda Wave Interferometry. *Sensors* **2020**, *20*, 4023.
15. Deroo, F.; Kim, J.Y.; Qu, J.; Sabra, K.; Jacobs, L.J. Detection of damage in concrete using diffuse ultrasound. *The Journal of the Acoustical Society of America* **2010**, *127*, 3315–3318.
16. Little, M.A. *Machine Learning for Signal Processing: Data Science, Algorithms, and Computational Statistics*; Oxford University Press, USA, 2019.
17. Bishop, C.M. *Pattern recognition and machine learning*; springer, 2006.
18. Lattanzi, D.; Miller, G.R. Robust automated concrete damage detection algorithms for field applications. *Journal of Computing in Civil Engineering* **2014**, *28*, 253–262.
19. Zhang, C.; Chang, C.c.; Jamshidi, M. Concrete bridge surface damage detection using a single-stage detector. *Computer-Aided Civil and Infrastructure Engineering* **2020**, *35*, 389–409.
20. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering* **2017**, *32*, 361–378.
21. Ariannejad, H. Numerical Simulation of Diffuse Ultrasonic Waves in Concrete **2019**.
22. Chen, G.; Pageot, D.; Legland, J.B.; Abraham, O.; Chekroun, M.; Tournat, V. Numerical modeling of ultrasonic coda wave interferometry in a multiple scattering medium with a localized nonlinear defect. *Wave Motion* **2017**, *72*, 228–243.
23. Wriggers, P.; Moftah, S.O. Mesoscale models for concrete: Homogenisation and damage behaviour. *Finite Elements in Analysis and Design* **2006**, *42*, 623–636.
24. Wriggers, P.; Löhnert, S. Characterization of concrete by a multiscale approach. International RILEM Conference on Material Science. RILEM Publications SARL, 2010, pp. 3–12.
25. Nitka, M.; Tejchman, J. Modelling of concrete behaviour in uniaxial compression and tension with DEM. *Granular Matter* **2015**, *17*, 145–164.

26. Saenger, E.H.; Gold, N.; Shapiro, S.A. Modeling the propagation of elastic waves using a modified finite-difference grid. *Wave Motion* **2000**, *31*, 77–92, [arXiv:1011.1669v3]. doi:10.1016/S0165-2125(99)00023-2.
27. Holla, V.; Vu, G.; Timothy, J.J.; Diewald, F.; Gehlen, C.; Meschke, G. Computational Generation of Virtual Concrete Mesostructures. *Preprints* **2021**, [https://doi.org/10.20944/preprints202106.0669.v1].
28. Van Mier, J.G. *Fracture processes of concrete*; Vol. 12, CRC press, 1996.
29. Concrete Mesostructure Generation using Python. https://pycmg.readthedocs.io/en/latest/.
30. Cundall, P.; Strack, A. A discrete numerical model for granular assemblies. *Geotechnique* **1979**, *29*, 47–65.
31. Šmilauer, V. Cohesive particle model using discrete element method on the Yade platform. PhD thesis, Czech Technical University, 2010.
32. Vu, G.; Iskhakov, T.; Timothy, J.J.; Schulte-Schrepping, C.; Breitenbücher, R.; Meschke, G. Cementitious Composites with High Compaction Potential: Modeling and Calibration. *Materials* **2020**, *13*, 4989.
33. Saenger, E.H.; Bohlen, T. Finite-difference modeling of viscoelastic and anisotropic wave propagation using the rotated staggered grid. *Geophysics* **2004**, *69*, 583–591.
34. Kocur, G.K.; Saenger, E.H.; Vogel, T. Elastic wave propagation in a segmented X-ray computed tomography model of a concrete specimen. *Construction and Building Materials* **2010**, *24*, 2393–2400.
35. Saenger, E.H. Time reverse characterization of sources in heterogeneous media. *NDT & E International* **2011**, *44*, 751–759.
36. Saenger, E.; Shapiro, S.; Keehm, Y. Seismic effects of viscous Biot-coupling: Finite difference simulations on micro-scale. *Geophysical Research Letters* **2005**, *32*.