12th International Conference on Fracture Mechanics of Concrete and Concrete Structures
FraMCoS-12

B.L.A. Pichler, Ch. Hellmich, P. Preinstorfer (Eds)

https://doi.org/10.21012/FC12.1109
MS06-2:2

USE OF CONVOLUTIONAL NEURAL NETWORKS FOR PREDICTING THE SHORT-TERM CREEP MODULUS OF CEMENT PASTE

MINFEI LIANG[*], YIDONG GAN[†], ERIK SCHLANGEN[*] AND BRANKO ŠAVIJA[†]

[*] Microlab, Faculty of Civil Engineering and Geosciences, Delft University of Technology
Stevinweg 1, 2628 CN Delft, the Netherlands
e-mail: minfei.liang@physics.ox.ac.uk; erik.schlangen@tudelft.nl; b.savija@tudelft.nl

[†] School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology
430074 Wuhan, China
e-mail: ygan@hust.edu.cn

Key words: cement paste; creep; convolutional neural network; machine learning

Abstract: Micromechanical and time dependent properties of cement paste can be predicted based on the microstructure by using analytical or numerical models. Herein, we propose an alternative approach for predicting the creep modulus of cement paste based on deep convolutional neural network (DCNN). The DCNN is trained using numerical simulation data obtained by the microscale lattice model, resulting in a database with more than 18000 samples. Then, 3 different DCNN architectures are built to learn from (part of) the database. Finally, the accuracy of DCNN prediction is tested on unseen samples. The proposed DCNN architectures can achieve excellent accuracy on the testing set, with the R^2 higher than 0.95. Furthermore, the distribution of creep moduli predicted by the DCNNs coincides with the original dataset. Further analyses of the feature maps show that the DCNNs can correctly capture the local importance of different microstructural phases on the predicted creep moduli. Therefore, it was concluded that a *well-trained* DCNN allows prediction of creep moduli based on microstructural images as input, which is computationally much more efficient compared to image segmentation and numerical simulation methods commonly used today. Of course, computational demands for training the network may be significant, but are needed in principle only once.

1 INTRODUCTION

In recent years, several investigations have been performed on the creep of hardened cement paste at the microscale, including the intrinsic creep properties of low- and high-density C-S-H [1]. Recently, the authors [2,3] proposed a method for characterizing the viscoelastic properties of hardened cement paste using micro-cantilever testing. Based on the experimental observations and the literature, the authors further developed am micro-scale lattice model for simulating the time dependent deformation of cement paste under constant load [4]. However, the model is relatively computationally expensive. Herein, we explore an alternative approach- the use of deep convolutional neural networks (DCNNs) – for predicting the viscoelastic properties of hardened cement paste.

DCNNs have been broadly used in the context of structural health monitoring and image segmentation of cracks in concrete. Herein, we use the previously developed microscale lattice model for creep prediction to create a database of creep strain curves based on 18,920 microstructures derived from X-ray computed tomography scans. Then, we train the DCNNs using this database. Finally, we evaluate the performance of the DCNNs

predictions when using unseen data.

2 MICROSCALE LATTICE CREEP MODEL

As with any data-driven approach, the performance of DCNN strongly depends on the size and the accuracy of training and testing data. Herein, we use the experimentally validated micro-scale lattice model to generate a database of creep moduli of cement pastes with different microstructure. First, 18,920 XCT images of the size 200x200 μm are segmented to generate the corresponding lattice meshes, which consist of 4 phases: high-density CSH (HD-CSH), low-density CSH (LD-CSH), unhydrated particles and pores. Then, viscoelastic deformation of the microstructure is calculated by the model, and the creep compliance curve is fitted according to power function law. Based on the curve, the creep modulus of different microstructures is derived, which is used to form the training database for the DCNNs.

To create a database of microstructures, XCT was used to scan 3 micro-cantilever beams of the size 1650x300x300 μm. From the scans, 9460 slices of 200x200 μm were cropped. The fabrication process of the micro-cantilever beams is described elsewhere [2,3]. Afterwards, each image is segmented into four phases (LD-CSH, HD-CSH, pores, unhydrated particles) according to the grayscale histogram, using the procedure described in [4].

The segmented microstructures provided the input for lattice modelling. Mapping of the segmented image onto a random lattice mesh is schematically shown in Figure.

The short-term creep behavior of cement paste is simulated under uniaxial compression using a 2D model for computational reasons. Local properties of different phases, including the tensile strength fi, elastic modulus E and creep modulus C are derived from the authors' previous testing results (Table 1).

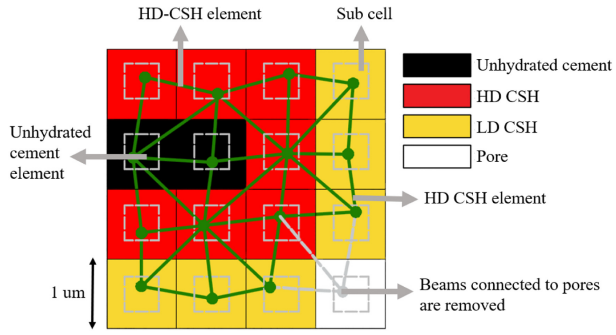


Figure 1: Mapping of an irregular lattice network onto a segmented microstructure.

Table 1: Local properties of different phases

Phase	f_t (MPa)	E (GPa)	C (GPa)
Uhydrated cement	614.7	84.2	-
LD CSH	52.2	21.3	670
HD CSH	82.8	26.4	990

With the boundary conditions and local material properties settled, an elastic calculation is firstly performed to get the initial mechanical response of the microstructure to the constant load. During the simulation, beam elements that meet the maximum stress criterion are removed from the mesh, as commonly done in the Delft lattice model and described elsewhere.

After the initial mechanical response of the microstructure to the immediate constant load is calculated, the creep calculations are conducted using an incremental local force method [4]. In each following step, a local force is calculated based on the local mechanical response of each node (calculated in first step) to mimic the short-term creep effect. First, the creep compliance function of CSH is expressed as:

$$C(t,t_0) = \frac{1}{E_c}\left(\frac{t-t_0}{\Delta t}\right)^{\beta} \qquad (1)$$

In which C(t, t0) is the creep compliance at time t when loaded at time t0; Δt is the time interval used in the calculation, which is 1 s in this study; Ec is the creep modulus of the CSH phases (Table 1); β is a constant which equals to 0.251 according to [5]. Then, based on the Euler's method, the incremental creep strain at each time step can be calculated as:

$$\varepsilon_{i+1}^{cr} - \varepsilon_i^{cr} = \frac{1}{E_c}\beta \Delta t \sigma_i \left(\frac{t - t_0}{\Delta t}\right)^{\beta - 1} \quad (2)$$

in which ε_i^{cr} is the creep strain at *i*-th step; σ_i is the local stress. Thereby, the local creep strain of each beam is dependent on the local stress distribution, which is converted to local force [4]. Based on the methods introduced above, the short-term creep behavior of a specific microstructure of cement paste can be calculated. For full implementation of the model, the reader is referred to our earlier work [4]. By fitting the calculated creep compliance curves with Eq. (1), the creep modulus can be derived. The fitted exponent β equals to 0.4285 and remains constant for all specimens. Therefore, we use only a single parameter (i.e., the creep modulus) to characterize the creep behavior of a specific cement paste microstructure.

3 DEEP CONVOLUTIONAL NEURAL NETWORK (DCNN)

Based on the microscale lattice creep model, the database for training the DCNNs can be produced. Data derived from 18,920 virtual specimens will be used to train a DCNN. To increase the number of data points, data augmentation is used.

By rotating the input microstructure by 90 degrees and running the lattice creep model with the same boundary condition, different creep behaviors will be derived. With the 9460 XCT images of different microstructure, this results in 18,920 sets of data in this database. Each data set includes a pair of an XCT image and a corresponding creep modulus.

3.1 DCNN architecture

Three kinds of DCNN architectures are tested: with 1 single convolutional layer, 2-stack consecutive convolutional layers, and 3-stack consecutive convolutional layers will be adopted, which are named as DCNN-1, DCNN-2 and DCNN-3, respectively. The DCNN is formulated by the following 6 types of layers: convolutional layer, max pooling layer, ReLU activation layer, global pooling layer, fully connected dense layer and linear activation layer. First, the input layer receives a grayscale XCT image representing a specific microstructure. Then, the image is passed through a series of consecutive convolutional layers and processed by a ReLU activation function and a max pooling layer. Subsequently, the stacked feature maps produced by the convolutional layers will be flattened by a global pooling layer and transferred into a 512*1 vector. The vector will then be passed through fully connected dense layers and finally output a scalar, which is the creep modulus that can describe the creep behavior of the input microstructure under axial compression. A schematic description and parameters of the DCNN's can be found in [6].

3.2 Training configuration, data preprocessing, and optimization

The training of DCNN can be conducted by importing the database into the designed network architecture. The three DCNN configurations will be compared in terms of performance. Before training starts, the database will be shuffled and grouped into 3 sets: a training set, a validation set and a testing set following the proportion ratio 6.4: 1.6: 2. Thus, there are 12,108, 3028 and 3784 samples in the training, validation and testing set, respectively. Then, the data of creep modulus in training set will be standardized by its max and min value to make sure the range of each sample point lie in the range of (0,1). Note that the standardization should only be conducted within the training set and cannot be extended to the testing set, since the statistical information of the testing set should be totally isolated from training set.

During the training process, the weights and biases of the filters and dense networks are updated to optimize the prediction performance on training set. Mean squared error is adopted as the loss function. Then, by a gradient descent algorithm, the weights and filters are updated to minimize the loss function through backpropagation of error. In this study, the Adam algorithm is used for optimization.

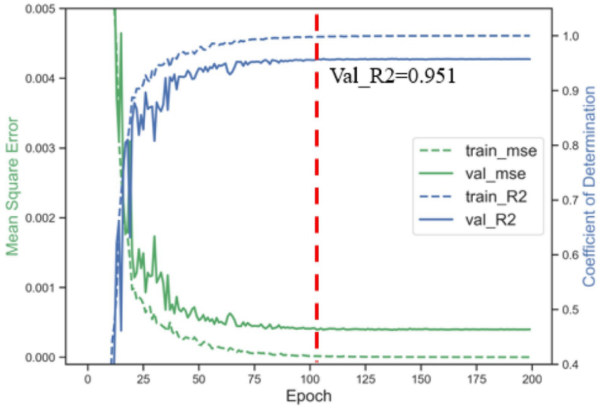
4 RESULTS AND DISCUSSION

4.1 Metrics and performance

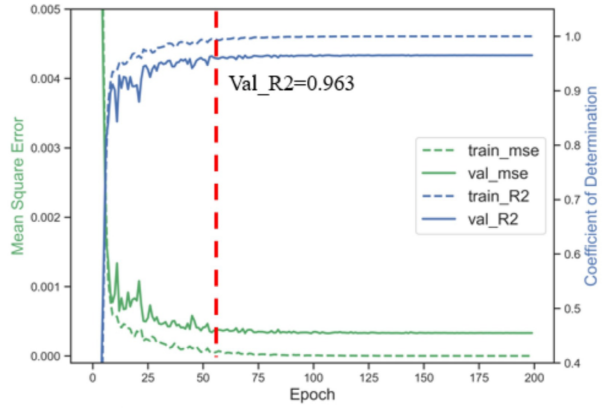
Three metrics are used to evaluate the performance of the DCNNs in training, validation and testing set, which are Mean Squared Error (MSE), Mean Absolute Error (MAE) and Coefficient of Determination (R^2). Their definitions can be found in general machine learning literature.

In the training process, MSE is adopted as the loss function that is used in gradient descent, and R^2 is used as the metrics for adjusting learning rate (see [6] for details). Therefore, the MSE and R^2 are used to describe the training history. In the training and validation set, the MSE and R2 of the 3 DCNNs are shown in Fig. 2 (a) ~ (c), and the learning rate history are shown in Fig. 2 (d). From Fig. 2, one can see that the 3 DCNNs all achieve high prediction performance in the training process. On the training and validation set, the R^2 of all the DCNNs are above 0.99 and 0.95, respectively, which indicates low risk of overfitting since the prediction performance of training and validation set are similar. Furthermore, the results also prove the adaptive adjustment of learning rate are effective for the training process of DCNNs: for DCNN-1, DCNN- 2 and DCNN-3, the learning rate decreases by 96.5%, 97.6%, and 97.8% in 200 epochs, respectively. However, comparing the difference of DCNNs, one can find the importance of having consecutive convolutional layers. With the 3-stack convolutional layers, DCNN-3 gives the best prediction while maintaining the best numerical stability since the first iterations. Besides, the fast decrease of learning rate of DCNN-3 also indicates that better numerical stability is achieved. Comparing DCNN-1 with the others, more distinguished effects of consecutive convolutional layers on both prediction performance and numerical stability can also be found.

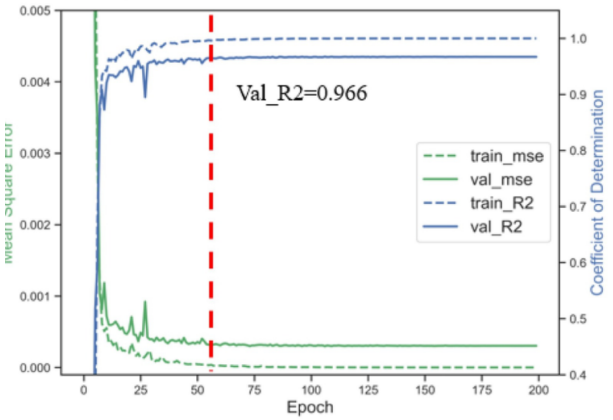
After the training process is finished, the DCNNs are tested by the 3784 data samples, which have been kept isolated since the beginning. The testing results of the DCNNs are shown in Fig. 3. The testing results of different DCNNs are consistent with that of the training and validation set: All the 3 DCNNs can achieve high accuracy over most samples within the unseen testing set. And more consecutive convolutional layers can improve the testing accuracy expressed by R^2, MSE and MAE.



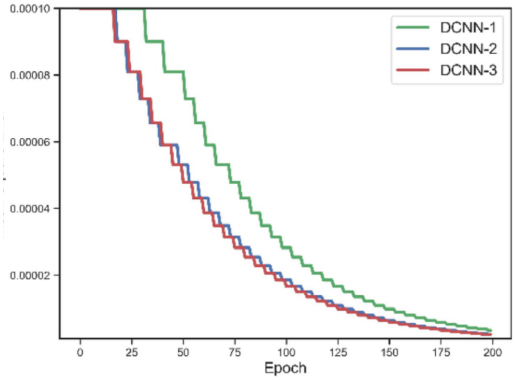
(a) Performance of DCNN-1



(b) Performance of DCNN-2

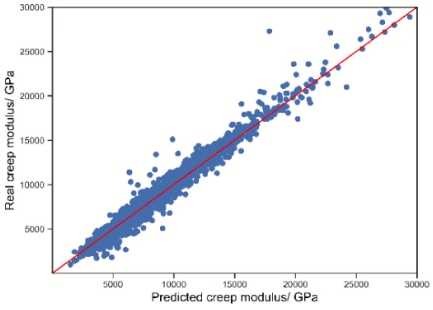


(c) Performance of DCNN-3

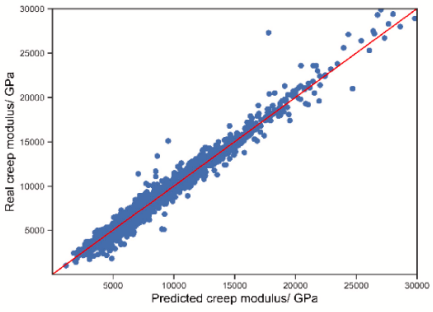


(d) Learning rate history

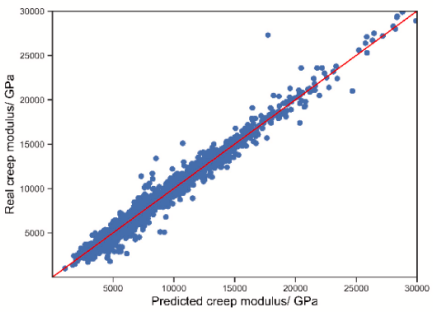
Figure 2: Training history



(a) DCNN-1, R^2 = 0.9678, MSE = 0.00391, MAE = 0.01285



(b) DCNN-2, R^2 = 0.9753, MSE = 0.00300, MAE = 0.01100



(c) DCNN-3, R^2 = 0.9757, MSE = 0.00295, MAE = 0.01088

Figure 3: Testing results

4.2 Correlation with microstructural parameters

To further validate the capabilities of the DCNN, the correlation between the microstructural parameters and prediction results in the testing set are compared. Simulation results showed the most significant microstructure parameters that influence the creep modulus are LD-CSH ratio, Porosity and Unhydrated Cement Ratio. Therefore, the correlation between these parameters and creep modulus predicted by the DCNNs are shown in Fig. 4. By comparison, the distribution of creep modulus predicted by the DCNNs highly coincides with the original distribution of Lattice modelling results in testing set.

It can also be shown that the DCNNs can capture the feature of an input microstructure and give its prediction. As mentioned in [6], the local features of the local feature of the microstructure are extracted by the filters to form a new input image block (a 3D matrix), which are called feature maps and are passed to the subsequent layers after ReLU activation and max pooling. Therefore, the feature maps indicate the local features identified by different convolutional layers. Taking DCNN-3 as an example, the feature maps produced at the end of every consecutive convolutional layer are shown in Fig. 5. Note that for each consecutive convolution layer, a stack of feature maps is produced. The results shown in Fig. 5 are the summation of these feature maps along its channel. The DCNNs keeps extracting the features and downscaling the size of feature maps. At the first layers, only the finer local features are extracted and compressed into smaller feature maps. While when it goes deeper into the DCNN, more details are compressed into smaller number of pixels and therefore such feature maps become highly abstract and uninterpretable.

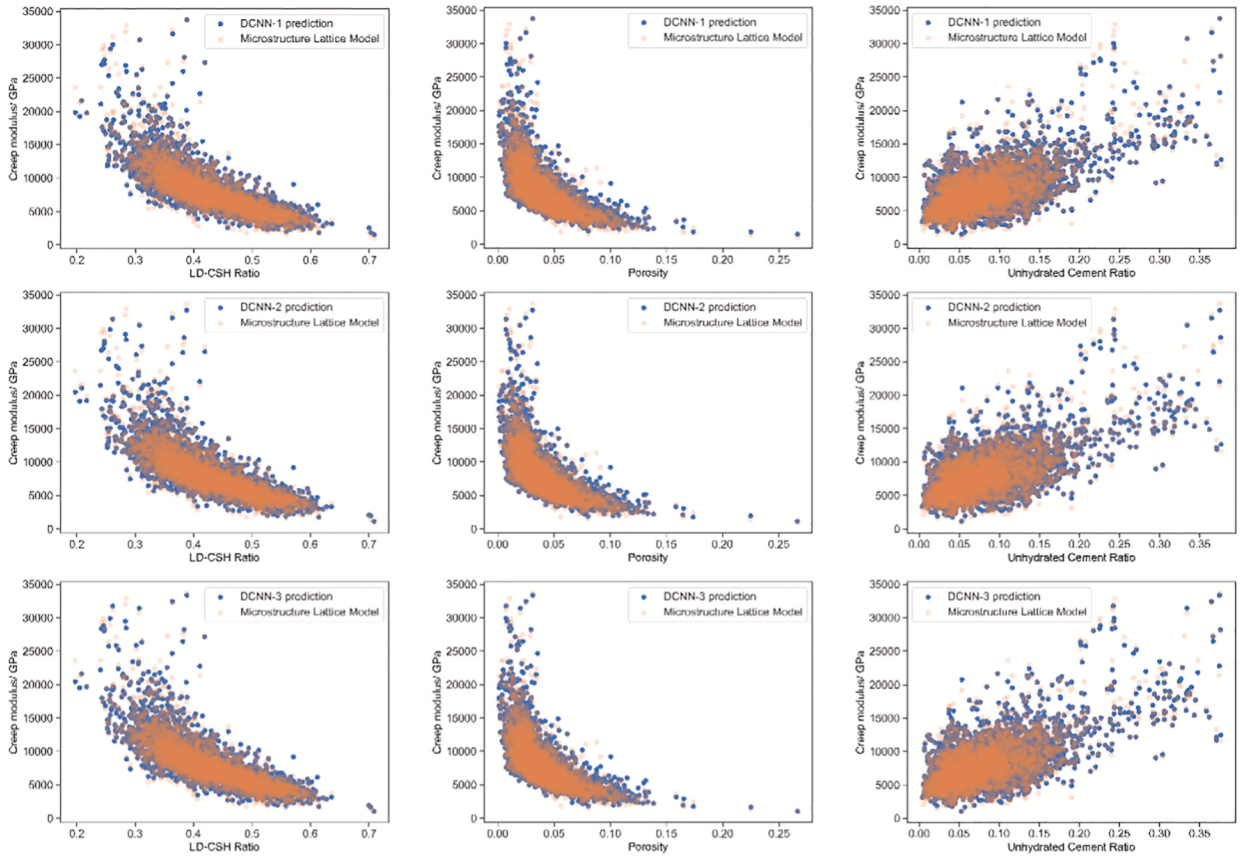


Figure 4: Correlation between predicted creep modulus and microstructural parameters

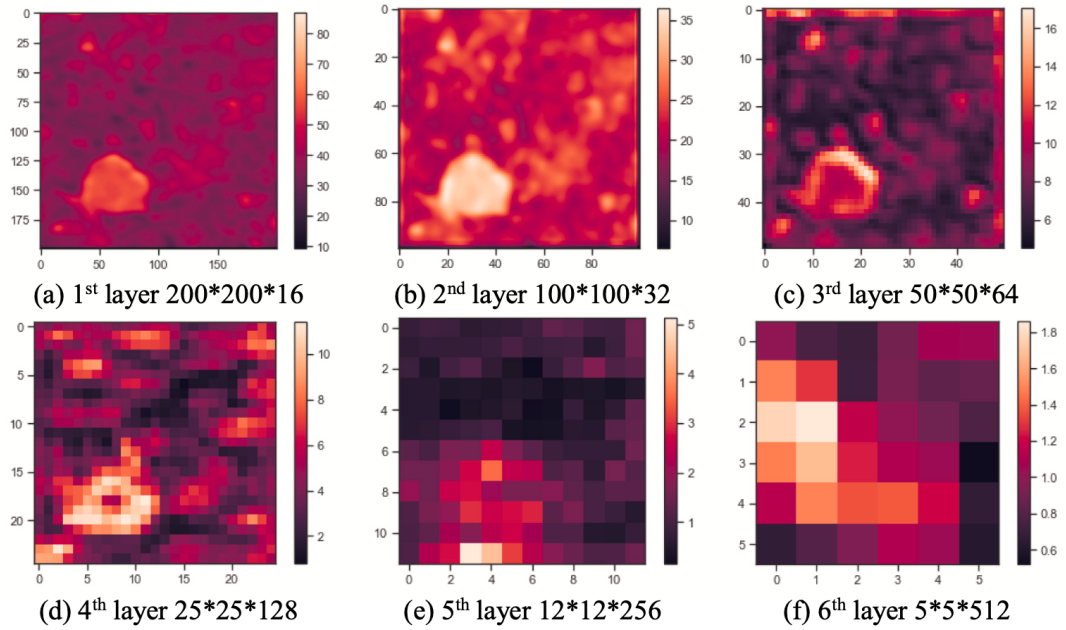


Figure 5: Feature maps extracted from DCNN-3

5 GENERAL DISCUSSION

We have shown that the DCNNs can accurately mimic the microscale lattice creep model. By comparison, the DCNNs save the computational resources for image segmentation and multiple incremental iterations of the microscale lattice creep model and directly predict the creep modulus with the input of a microstructure represented by a raw XCT image. The capabilities of extracting and integrating local features from highly heterogeneous microstructures make DCNN a powerful tool in predicting the short-term creep of cement paste. Moreover, such capabilities indicate strong potential of DCNN in other computationally- intensive tasks for homogenization of heterogenous materials.

However, it must be emphasized that the performance of the DCNNs highly depends on the microscale lattice creep model, which generated the datasets for the whole training process. The applications of DCNNs of this study is limited to the scenario of hardened four-phase cement paste under immediate axial compression. If any other phases (e.g., CH) of cement paste play an essential role, or other boundary conditions are to be considered, the microscale lattice creep model will need to be adjusted first. After new datasets that reflect the influence of new phases or different boundaries are generated, the DCNNs can be retrained to gain accuracy on the new tasks. If the changes of the microscale lattice creep model are minor, then transfer learning can be applied in this process. In this case, only parts of the DCNN network must be adjusted to gain good accuracy for new datasets, which not only saves computational resources in training process but also tends to gain higher accuracy.

6 CONCLUSIONS

In this study, we built a database that contains 18,920 microstructures and their corresponding creep moduli using an experimentally- validated microscale lattice model for short-term creep. Then, DCNNs with different numbers of consecutive convolutional layers are established to predict the creep modulus given the microstructure as input. Finally, the distribution of creep modulus predicted by DCNNs and the lattice model are compared, and the local importance of microstructure are analyzed based on the output of feature maps. Through this study, the following conclusions can be drawn:

- The modelling results of 18,920 microstructures show that the adopted microscale lattice model captures well the correlations between the creep behaviors and microstructural parameters,
- Due to the amplification of receptive field, the consecutive convolution layers can promote the prediction accuracy of DCNNs and meanwhile maintain both the numerical stability.
- All the DCNNs adopted in this study can achieve high accuracy in predicting the creep modulus of unseen microstructure, with the values of R^2 all above 0.96. Moreover, the correlation between creep modulus predicted by the DCNNs and microstructural parameters is consistent with that of the original database.

ACKNOWLEDGEMENTS

Branko Šavija acknowledges the financial support from the European Research Council (ERC) within the framework of the ERC Starting Grant Project “Auxetic Cementitious Composites by 3D printing (ACC-3D)”, Grant Agreement Number 101041342. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

[1] Wei, Y., Liang, S., and Gao, X. 2017. Indentation creep of cementitious materials: experimental investigation from nano to micro-length scales. *Con. Build.*

Mater. **143**:222-233

[2] Gan, Y., Vandamme, M., Zhang, H., Chen, Y., Schlangen, E., van Breugel, K., and Šavija, B. 2020. Micro-cantilever testing on the short-term creep behaviour of cement paste at micro-scale. *Cem. Concr. Res.* **134:**106105

[3] Gan, Y., Vandamme, M., Chen, Y., Schlangen, E., van Breugel, K., and Šavija, B. 2021. Experimental investigation of the short-term creep recovery of hardened cement paste at micrometre length scale. *Cem. Concr. Res.* **149**:106652

[4] Gan, Y., Romero Rodriguez, C., Zhang, H., Schlangen, E., van Breugel, K., and Šavija, B. 2021. Modeling of microstructural effects on the creep of hardened cement paste using an experimentally informed lattice model. *Comput. Aided Civ. Infrastruct. Eng.* **36**(5):560-576

[5] Königsberger, M., Irfan-ul-Hassan, M., Pichler, B., and Hellmich, C. 2016. Downscaling based identification of nonaging power-law creep of cement hydrates. *J. Eng. Mech.* **142**(12):04016106

[6] Liang, M., Gan, Y., Chang, Z., Wan, Z., Schlangen, E., and Šavija, B. 2022. Microstructure-informed deep convolutional neural network for predicting short-term creep modulus of cement paste. *Cem. Concr. Res.* **152**:106681