Deep neural network method for predicting the mechanical properties of composites

Cite as: Appl. Phys. Lett. **115**, 161901 (2019); doi: 10.1063/1.5124529
Submitted: 15 August 2019 · Accepted: 30 September 2019 ·
Published Online: 14 October 2019

View Online Export Citation CrossMark

Sang Ye, Bo Li, Qunyang Li, Hong-Ping Zhao, and Xi-Qiao Feng[a)]

AFFILIATIONS

Institute of Biomechanics and Medical Engineering, AML, Department of Engineering Mechanics, Tsinghua University, Beijing 100084, China

[a)]Author to whom correspondence should be addressed: fengxq@tsinghua.edu.cn

ABSTRACT

Determining the macroscopic mechanical properties of composites with complex microstructures is a key issue in many of their applications. In this Letter, a machine learning-based approach is proposed to predict the effective elastic properties of composites with arbitrary shapes and distributions of inclusions. Using several data sets generated from the finite element method, a convolutional neural network method is developed to predict the effective Young's modulus and Poisson's ratio of composites directly from a window of their microstructural image. Through numerical experiments, we demonstrate that the trained network can efficiently provide an accurate mapping between the effective mechanical property and the microstructures of composites with complex structures. This study paves a way for characterizing heterogeneous materials in big data-driven material design.

Determining the mechanical properties of composites is of great significance for their applications in a wide diversity of engineering fields.[1–3] Besides experimental characterization,[4,5] some theoretical homogenization methods have been established to estimate the macroscopically average properties of composites, e.g., self-consistent method, Mori–Tanaka's mean-field method,[6] and Chamis's micromechanical model.[7] Most of these are based on Eshelby's solution of inclusions with elliptical or other regular shapes[8] and thus hard to predict the macroscopically effective properties of composites with complex microstructures. For such complex composites, the finite element method (FEM) is normally utilized to calculate the material properties by defining a representative volume element (RVE).[3,9,10] The RVE should be sufficiently large in order to rule out the effect of RVE sizes, and, therefore, the FEM calculation is usually tedious and highly time-consuming. In addition, the RVE method is specifically sensitive to the geometric and topological features of the microstructures of composites.

The mechanical properties of a composite depend on its manufacturing process, constituent phases, and microstructures, making their determination a complex nonlinear and material-specific problem.[10] Recently, the artificial intelligence method has been extended to various problems of materials science and engineering, e.g., predicting their crystal structures and stability,[11,12] atomic and molecular properties,[13–16] complex phase transitions,[17,18] and service life.[19] In particular, the deep neural network (DNN) method has been proven to be efficient in discovering meaningful structures in data. With a sufficient amount of training data and a specific target, this method is capable of extracting the high-dimensional feature vector from the original data and learning the nonlinear mapping from the feature vector to a desired output.[11,20]

In this Letter, we propose a DNN method to predict the mechanical properties of composites with complex microstructures. Through the implementation of RVE homogenization in the FEM, we first generate multiple data sets, which contain pairs of images describing the microstructures, the properties of constituents, and the mechanical properties of the composites. With these data sets, we train a DNN model to predict the mechanical properties of a composite from a window of its microstructural image. We demonstrate that this DNN model can accurately and efficiently predict the mechanical properties of composites with various complex microstructures, which are difficult to predict with the traditional homogenization methods that use the Eshelby inclusion solution.

In the micromechanics of composites, the RVE concept is widely adopted to estimate the macroscopically effective mechanical responses.[21,22] An RVE should be sufficiently large to avoid significant dependence of the numerical results on its sizes.[23] Besides, periodic or other appropriately specified boundary conditions should be applied to the RVE. Under loading, the total elastic strain energy U stored in the RVE is expressed by

$$U = \frac{1}{2}\bar{\sigma}_{ij}\bar{\varepsilon}_{ij}V_{\mathrm{RVE}}, \tag{1}$$

where $\bar{\varepsilon}_{ij}$ and $\bar{\sigma}_{ij}$ are the macroscopically averaged strains and stresses, respectively, and V_{RVE} is the volume of the RVE.[24] Hereafter, the subscripts i, j, k, and l take values 1–3, and Einstein summation convention is adopted for repeated indexes. U can also be calculated from the microscopic fields by

$$U = \frac{1}{2}\int_{V_{\mathrm{RVE}}} \sigma_{ij}(\boldsymbol{x})\varepsilon_{ij}(\boldsymbol{x})\mathrm{d}V, \tag{2}$$

where $\sigma_{ij}(\boldsymbol{x})$ and $\varepsilon_{ij}(\boldsymbol{x})$ denote the microscopic stresses and strains at position $\boldsymbol{x}$ in the RVE, respectively. Then, the elastic stiffness components C_{ijkl} can be determined from the calculation results of RVE by

$$\bar{\sigma}_{ij} = C_{ijkl}\bar{\varepsilon}_{kl}. \tag{3}$$

Assume all material components to be linear elastic, and we consider only the two-dimensional case in this Letter though the idea presented in this work can be extended to three-dimensional problems. For a plane-stress problem, Young's modulus (E) and Poisson's ratio (ν) of the composite can be derived as

$$E = \frac{C_{1111} + C_{2222}}{2}(1-\nu^2), \quad \nu = \frac{2C_{1122}}{C_{1111} + C_{2222}}. \tag{4}$$

In our study, periodic boundary conditions (PBCs) are applied to the RVE, which can be expressed as[25]

$$u_i^{\mathrm{b}+} - u_i^{\mathrm{b}-} = \bar{\varepsilon}_{ik}\Delta x_k^{\mathrm{b}}, \tag{5}$$

where $u_i^{\mathrm{b}+}$ and $u_i^{\mathrm{b}-}$ are the displacements on a pair of nodes $x_k^{\mathrm{b}+}$ and $x_k^{\mathrm{b}-}$ of two opposite boundary surfaces, and $\Delta x_k^{\mathrm{b}} = x_k^{\mathrm{b}+} - x_k^{\mathrm{b}-}$, with superscript b indicating a quantity pertaining to boundary. We combine the FEM analysis with the RVE concept to account for the geometric features of complex microstructures which are hard to be modeled as elliptical inclusions.[26,27] We develop a Python script that can automatically compute the effective mechanical properties of the composite from a file that contains primary RVE information. Our Python script has been verified by using both two- and three-dimensional models and comparing with previous results in the literature.[10,25] For the sake of simplicity, we assume that all interfaces between the inclusions and the matrix are perfectly bonded, though the proposed method can be extended to account for the effects of imperfect interfaces.

In our DNN method, we first generate a large number of RVEs with various types of complex structures, consisting of arbitrary, either regular or irregular, shapes of inclusions. The microstructural images of composites, and the mechanical properties of the constituent materials as well, are stored in the database. We then simulate all different RVE structures of composites using the Python script to attain their effective mechanical properties, which will be treated as the ground truth in the development of our DNN prediction approach.

For the purpose of obtaining the effective mechanical properties from a window of microstructural images of composites, we use a specific type of deep learning architecture, i.e., convolutional neural network (ConvNet),[28–30] which has been widely used in the field of image processing. The ConvNet consists of a sequence of computationally nonlinear layers, which are able to gradually extract representations of images with higher-level abstractions.[31] A schematic representation of the ConvNet used in our approach is shown in Fig. 1. It has five convolutional layers connected by four fully connected layers. The convolution kernels are learned in a hierarchical manner, which is composed of low-level features to generate more complex patterns.[32] The gray image corresponding to a composite is input at the beginning of the ConvNet, and then the corresponding mechanical properties (Young's modulus and Poisson's ratio) are derived at the output layer of the ConvNet (Fig. 1).

Considering the huge difference between the values of elastic modulus E ($\sim 10^{11}$ Pa) and Poisson's ratio ν($\sim$0.3), and the efficiency during training and prediction, we use different activation functions in the output layer of the ConvNet. For Young's modulus, we choose the rectified linear unit (ReLU) function[33]

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}, \tag{6}$$

while the following variant form of Sigmoid function is used for Poisson's ratio:

$$f(x) = \frac{1}{2}\sigma(x) = \frac{1}{2(1+e^{-x})}. \tag{7}$$

The proposed ConvNet can automatically learn valid representations describing the geometric configurations and discard those

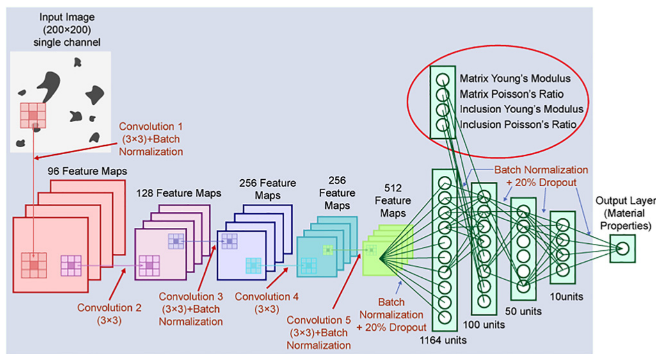

FIG. 1. Schematic of the convolutional neural network (ConvNet) used for prediction of the effective mechanical properties of composites. A 200 × 200 gray image, describing the microstructures of a composite, is presented as the input of the ConvNet. A learnable kernel is convolved across the input image, and the scalar product between the kernel and the input at every position is computed. This results in a two-dimensional map (in red) of that kernel at each spatial position, which is then added by a learnable bias of the same size. The input image is convolved with total 96 different 1st layer kernels, each of which has the sizes of 3 × 3, using a stride of 2 in both x and y directions. The resulting feature maps are then passed through a rectified linear unit (ReLU) function (not shown here). Similar operations are repeated in layers 2, 3, 4, and 5, except that batch normalization is performed after layers 1, 3, and 5. The last four layers are fully connected. The first layer takes the features from the top convolutional layer as the input in the vector form and is multiplied by a matrix of learnable weights to obtain a vector output, which is passed to the next fully connected layer as input. Dropout and batch normalization are used after every fully connected layer. The final layer is a ReLU function for Young's modulus and a variant of Sigmoid function for Poisson's rate. Minimizing the prediction error, the above-mentioned kernels are learned and will be activated when a similar feature appears in the input.[11] The red ellipse, which contains four input nodes to accept the mechanical properties of the constituent phases, is added to the ConvNet in the data set No. 3.

unimportant details for those parameters to be determined. To this end, we use a loss function to quantify the difference between the effective mechanical properties of the ConvNet's prediction and the ground truth and then adjust the parameters (i.e., kernels and biases in convolutional layers and weights in fully connected layers described in Fig. 1) of the ConvNet through backpropagation.[34,35] Iterative adjustments are made using a large amount of data to minimize the loss function. Since predicting the mechanical properties is a regression problem, we choose the mean square error (MSE) as the objective function, which is denoted as

$$L[Y, f(X)] = \frac{1}{n}\sum_{i=1}^{n}[Y - f(X)]^2, \tag{8}$$

where n is the batch size, X is the input gray image representing the microstructures of the composite, Y is the effective mechanical property that can be treated as ground truth of the composite, which could be Young's modulus, Poisson's ratio, or some other mechanical parameters, and $f(X)$ denotes the corresponding prediction of the ConvNet. The mechanical parameters can also be treated in the log space when they vary over a wide range.

During every epoch of training, we save the trained model and perform verification to prevent overfitting. We will select the model without overfitting and not adopt the early stopping technique. To improve the accuracy of the ConvNet's prediction, we use dropout[36] in three fully connected layers. Batch normalization[37] in several convolutional and fully connected layers is also introduced to mitigate the effects of initialization and accelerate the training of the deep neural network.

Next, we consider a composite consisting of a matrix and arbitrarily shaped inclusions. The elastic properties of the two phases are taken as $E^{\mathrm{Mat}} = 68.3$ GPa and $\nu^{\mathrm{Mat}} = 0.3$ for the matrix, and $E^{\mathrm{Inc}} = 379.3$ GPa and $\nu^{\mathrm{Inc}} = 0.1$ for the inclusion. A data set (No. 1) comprising 20 000 files describing the geometric configurations of RVEs is randomly generated. Each file represents a composite containing inclusions with arbitrary numbers, sizes, and shapes. Three examples are shown in Fig. 2(a). For every sample, the effective Young's modulus and Poisson's ratio are determined by using the Python script described above. Then, these data sets are randomly divided into training, validation, and test parts in the ratio of 3:1:1 (the same participation is applied for all data sets used). Using 60 repeated epochs of training, the root mean square errors (RMSEs) yielded by the ConvNet (without the red ellipse part in Fig. 1) in the verification set are 0.38 GPa and 0.0010 for Young's modulus and Poisson's ratio, respectively. The effective mechanical properties for the test set predicted by the ConvNet are plotted against those generated by the RVE homogenization method in Figs. 2(b) and 2(c). Those points are concentrated in the vicinity of the dashed line $y = x$, indicating that the proposed DNN model can perfectly learn the nonlinear mapping between the image describing the complex microstructures and the effective mechanical properties. The results also demonstrate that the DNN model works well also for other material systems that have not been encountered in the preceding training process.

The proposed DNN approach allows to handle composites that contain a large number of various types of inclusions with different properties. To demonstrate this, we construct a similar data set (No. 2) consisting of a matrix and two different reinforcing phases. The elastic

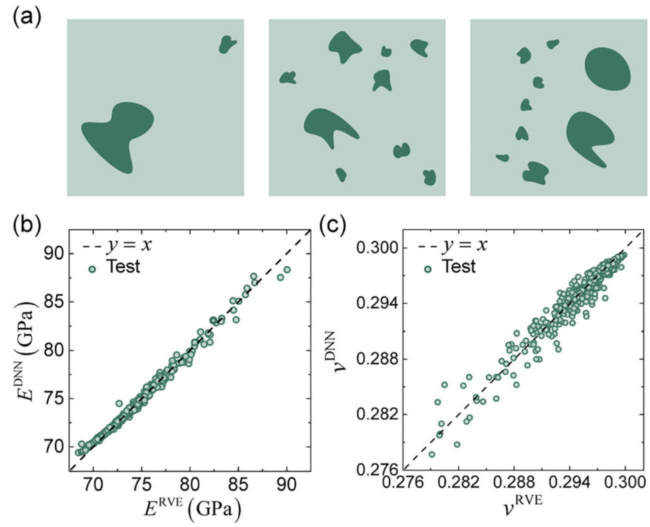

FIG. 2. DNN prediction of the mechanical properties of composites consisting of two different components. (a) Typical examples of complex microstructures. The effective Young's modulus E^{DNN} (b) and Poisson's ratio ν^{DNN} (c) predicted by the ConvNet are plotted against those generated by the RVE homogenization method. E^{RVE} and ν^{RVE} are the effective Young's modulus and Poisson's ratio generated by the RVE method, respectively. The black lines (dashed) are the identity lines serving as the references.

properties of the matrix are taken as above, and the elastic parameters of inclusions are $E_1^{\mathrm{Inc}} = 379.3$ GPa and $\nu_1^{\mathrm{Inc}} = 0.1$ for the first inclusion phase, and $E_2^{\mathrm{Inc}} = 281.5$ GPa and $\nu_2^{\mathrm{Inc}} = 0.4$ for the second inclusion phase. The number, sizes, and shapes of inclusions differ from each other. We use the ConvNet with the same structure as the data set No. 1 for training and calculate the relative error on the test set by

$$\frac{|P^{\mathrm{DNN}} - P^{\mathrm{RVE}}|}{P^{\mathrm{RVE}}} \times 100\%, \tag{9}$$

where P stands for the considered mechanical parameter such as Young's modulus E and Poisson's ratio ν. The distributions of the relative errors are plotted in Fig. 3(a). It shows that the ConvNet has a high accuracy in predicting the mechanical properties of composites composed of multiple phases, with a relative error lower than 4.0%.

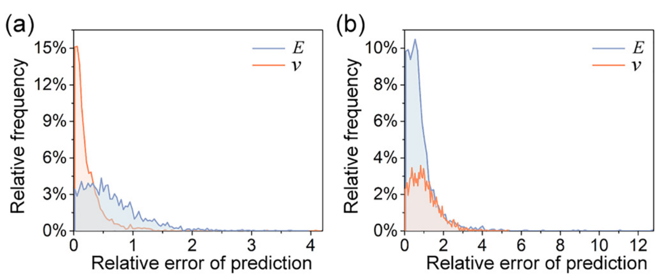

FIG. 3. The relative error frequency distribution curve of the obtained mechanical properties. (a) Composites composed of three phases. (b) Composites composed of two phases where the mechanical properties of the constituent materials can vary within a wide range.

In the previous two data sets, we only change the geometric configurations of the RVE but maintain the elastic properties of the constituent materials. To further investigate whether the ConvNet can fully learn the nonlinear mapping of mechanical properties of composites with complex microstructures, we generate a new data set (No. 3) containing 40 000 data, where the composites consist of a matrix reinforced by dispersed inclusions with elastic properties varying within $E^{\rm Mat} \in [60\,{\rm GPa},\ 100\,{\rm GPa}]$ and $\nu^{\rm Mat} \in [0.3,\ 0.4]$ for the matrix, and $E^{\rm Inc} \in [370\,{\rm GPa},\ 410\,{\rm GPa}]$ and $\nu^{\rm Inc} \in [0.1,\ 0.2]$ for the inclusions. Minor modifications are made to the convolutional network in this case. Four nodes are added to the first fully connected layer for the elastic property input of the two components (the red ellipse part in Fig. 1). After 80 epochs of training the modified ConvNet, the relative errors of prediction are shown in Fig. 3(b). Among the 8000 calculations in the test set, there exist a few samples where the relative errors of the obtained Young's modulus may reach 12%, but almost all samples have a relative error lower than 4%. The predicted Poisson's ratio is also sufficiently accurate, with the relative errors smaller than 3%. Therefore, our DNN method can efficiently and robustly grasp the mapping from the microstructural features and elastic properties of the constituent phases to the macroscopically effective mechanical properties of the composites.

In summary, we have proposed a ConvNet, on the basis of the DNN method, to tackle the challenge of predicting the macroscopically effective properties of composites with complex microstructures. From a window of geometric configuration image, our ConvNet can accurately and efficiently predict the mechanical properties of composites containing complex inclusions with arbitrary number, sizes, shapes, and properties. This method is easy to use and holds broad application prospects in big data-driven material design and nondestructive testing of structures with defects or damage. It should be noticed that only two-dimensional problems have been addressed as an illustration of this technique in this Letter. Considering that three-dimensional convolution methods are available in many popular machine learning frameworks, our method can be extended easily to predict the mechanical properties of three-dimensional composites[38] by invoking a three-dimensional variant of ConvNet.[39–41] In addition, the proposed methodology can also be applied to calculate other physical properties of composite materials, e.g., coefficients of heat conduction and thermal expansion, fatigue life, strength, and toughness.

Support from the National Natural Science Foundation of China (Grant Nos. 11432008, 11921002, 11620101001, and 11672161) is acknowledged.

REFERENCES

[1] W. Ogierman and G. Kokot, Mechanics 21, 118 (2015).
[2] O. Sen, S. Davis, G. Jacobs, and H. S. Udaykumar, J. Comput. Phys. 294, 585 (2015).
[3] X. Y. Zhou, P. D. Gosling, C. J. Pearce, Z. Ullah, and L. Kaczmarczyk, Int. J. Solids Struct. 80, 368 (2016).
[4] S. Ujihashi, Compos. Struct. 23, 149 (1993).
[5] G. Tagliavia, M. Porfiri, and N. Gupta, J. Compos. Mater. 43, 561 (2009).
[6] T. Mori and K. Tanaka, Acta Metall. 21, 571 (1973).
[7] C. C. Chamis, SAMPE Q. 15, 14 (1984).
[8] S.-Z. Lin, L.-Y. Zhang, J.-Y. Sheng, B. Li, and X.-Q. Feng, Arch. Appl. Mech. 86, 219 (2016).
[9] G. D. Cheng, Y. W. Cai, and L. Xu, Acta Mech. Sin. 29, 550 (2013).
[10] S. L. Omairey, P. D. Dunning, and S. Sriramula, Eng. Comput. 35, 567 (2019).
[11] A. Ziletti, D. Kumar, M. Scheffler, and L. M. Ghiringhelli, Nat. Commun. 9, 2775 (2018).
[12] W. Ye, C. Chen, Z. Wang, I.-H. Chu, and S. P. Ong, Nat. Commun. 9, 3800 (2018).
[13] Q. Zhou, P. Tang, S. Liu, J. Pan, Q. Yan, and S. C. Zhang, Proc. Natl. Acad. Sci. U.S.A. 115, E6411 (2018).
[14] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, Phys. Rev. X 8, 041048 (2018).
[15] X. Z. Yang, X. Liu, A. Q. Zhang, D. W. Lu, G. Li, Q. H. Zhang, Q. Liu, and G. B. Jiang, Nat. Commun. 10, 1620 (2019).
[16] H. Chan, M. J. Cherukara, B. Narayanan, T. D. Loeffler, C. Benmore, S. K. Gray, and S. K. R. S. Sankaranarayanan, Nat. Commun. 10, 379 (2019).
[17] Q. Zhu, A. Samanta, B. X. Li, R. E. Rudd, and T. Frolov, Nat. Commun. 9, 467 (2018).
[18] S. Scandolo, Proc. Natl. Acad. Sci. U.S.A. 116, 10204 (2019).
[19] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis *et al.*, Nat. Energy 4, 383 (2019).
[20] H. Salmenjoki, M. J. Alava, and L. Laurson, Nat. Commun. 9, 5307 (2018).
[21] W. J. Drugan and J. R. Willis, J. Mech. Phys. Solids 44, 497 (1996).
[22] R. Hill, J. Mech. Phys. Solids 11, 357 (1963).
[23] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures* (American Mathematical Society, 2011).
[24] C. T. Sun and R. S. Vaidya, Compos. Sci. Technol. 56, 171 (1996).
[25] F. Ye and H. Wang, e-print arXiv:1703.03930.
[26] Z. Xia, F. Ju, and K. Sasaki, Finite Elem. Anal. Des. 43, 649 (2007).
[27] K. Xu and X. W. Xu, Mater. Sci. Eng. A 487, 499 (2008).
[28] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, Neural Comput. 1, 541 (1989).
[29] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, Proc. IEEE 86, 2278 (1998).
[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, in *Advances in Neural Information Processing Systems 25*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc., 2012), p. 1097.
[31] M. D. Zeiler and R. Fergus, in *Computer Vision-ECCV 2014*, edited by D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars (Springer, 2014), Part I, Vol. 8689, p. 818.
[32] Editor, Nat. Biomed. Eng. 1, 0036 (2017).
[33] V. Nair and G. E. Hinton, paper presented at Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010.
[34] D. P. Kingma and J. Ba, e-print arXiv:1412.6980.
[35] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Nature 323, 533 (1986).
[36] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, J. Mach. Learn. Res. 15, 1929 (2014).
[37] S. Ioffe and C. Szegedy, e-print arXiv:1502.03167.
[38] D. L. Shi, X. Q. Feng, Y. G. Y. Huang, K. C. Hwang, and H. J. Gao, J. Eng. Mater. Technol.-Trans. ASME 126, 250 (2004).
[39] T. Du, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, in Proceedings of the IEEE International Conference on Computer Vision (2015), p. 4489.
[40] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, in Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), LNCS 9901, edited by S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells (Springer, 2016), p. 424.
[41] F. Milletari, N. Navab, and S.-A. Ahmadi, in Proceedings of the Fourth International Conference on 3D Vision (2016), p. 565.