ICMA 2021

1st International Conference on Micromachines and Applications

15–30 APRIL 2021 | ONLINE

A deep learning-based approach to uncertainty quantification for polysilicon MEMS

micromachines MDPI

José Pablo Quesada-Molina[1,2,*], Stefano Mariani[1]

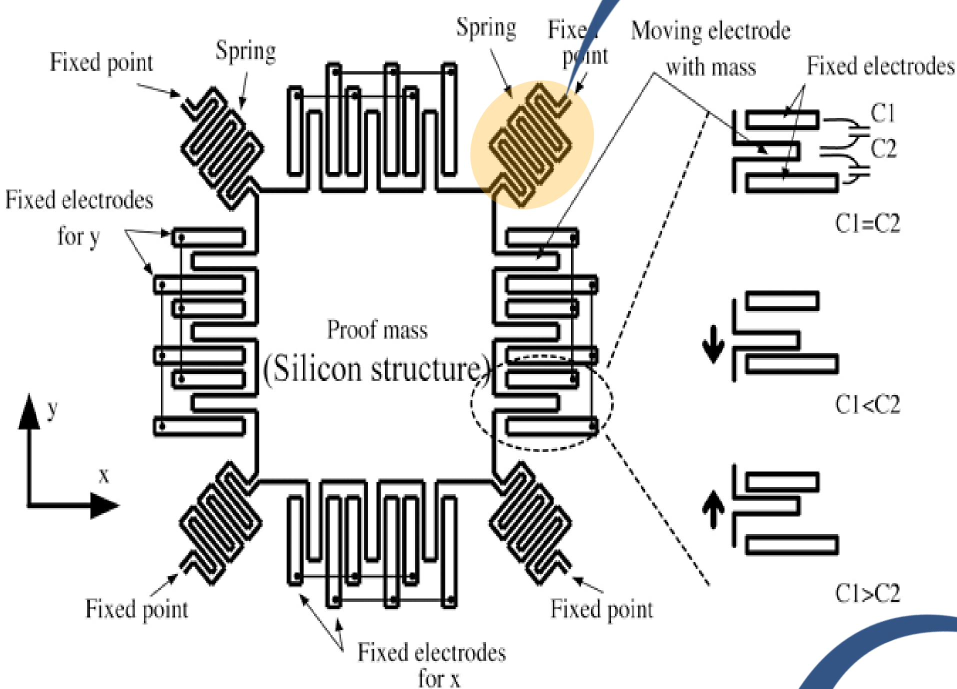
[1] Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano (Italy)
[2] Department of Mechanical Engineering, University of Costa Rica, San Pedro Montes de Oca, San José (Costa Rica)

* Corresponding author: josepablo.quesada@mail.polimi.it

POLITECNICO MILANO 1863 UNIVERSIDAD DE COSTA RICA

Graphical Abstract

1. Typical structure of an inertial MEMS

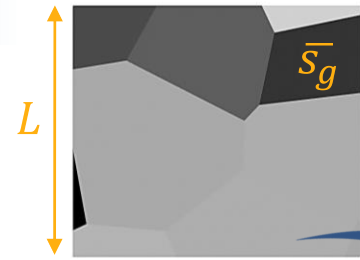


2-axis acceleration sensor

Reference [1]

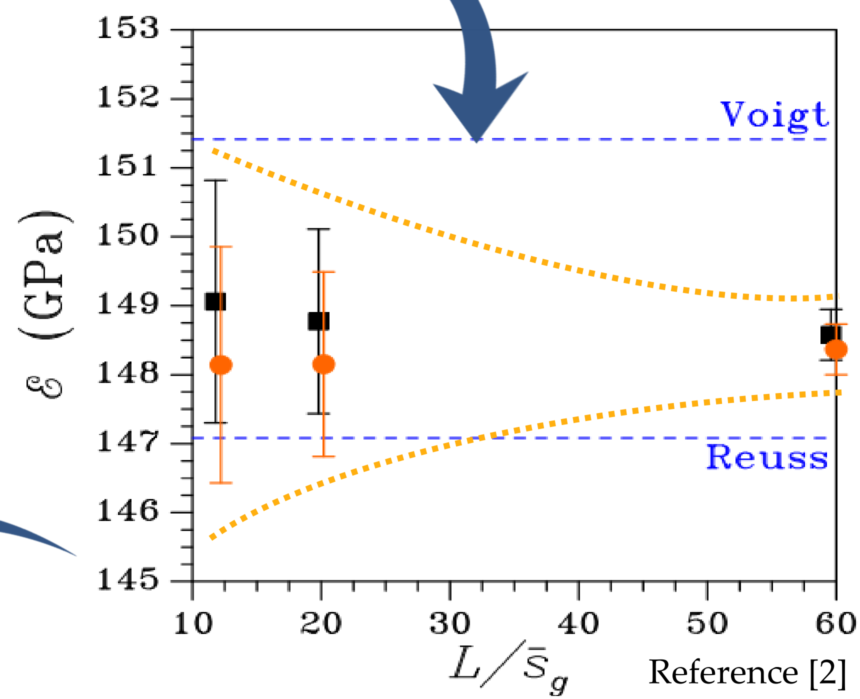
SVE

Statistical Volume Element



2. Overall properties are realization-dependent

Length scale separation principle for RVE does not hold ($L/\bar{s}_g \neq \infty$)



Reference [2]

3. Scattering in the overall properties

Dependent of the micromechanical features of the polycrystal: topology of grain boundary and lattice orientations

4. Idea: Training a neural network to catch this variation automatically !

ICMA 2021

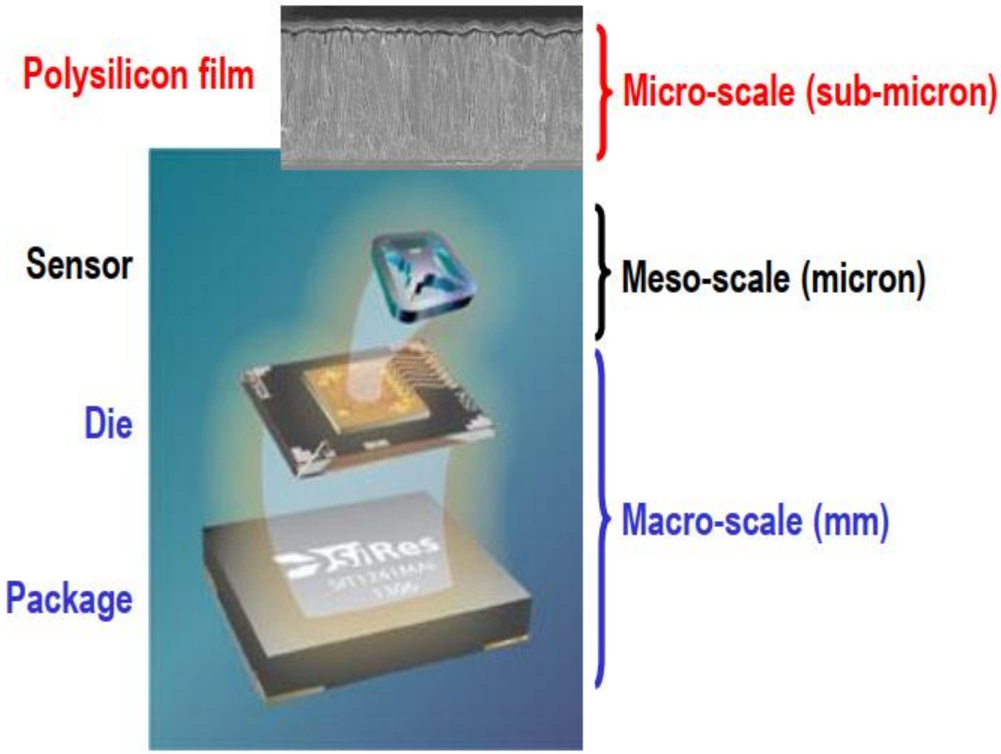
Abstract

The path towards miniaturization for micro electro-mechanical systems (MEMS) has recently increased the effects of stochastic variability at the (sub)micron scale on the overall performance of the devices. We recently proposed and designed an on-chip testing device to characterize two sources of variability that majorly affect the scattering in the response to the external actions of inertial (statically determinate) micromachines: the morphology of the polysilicon film constituting the movable parts of the device; and the environment-affected overetch linked to the microfabrication process. A fully stochastic model of the entire device has been set to account for these two sources on the measurable response of the devices, e.g. in terms of the relevant C-V curves up to pull-in. A complexity in the mentioned model is represented by the need to assess the stochastic (local) stiffness of polysilicon, depending on its unknown (local) microstructure. In this work, we discuss a deep learning approach to the micromechanical characterization of polysilicon films, based on artificial neural networks (NNs). Such NNs extract relevant features of the polysilicon morphology from SEM-like Voronoi tessellation-based digital microstructures. The NN-based model or surrogate is shown to correctly catch size effects at a varying ratio between the characteristic size of the structural components of the device, and the morphology-induced length scale of the aggregate of silicon grains. This property of the model looks indeed necessary, to prove the generalization capability of the learning process, and to next feed Monte Carlo simulations resting on the model of the entire device.

Keywords: Polysilicon MEMS; stochastic variability; homogenization; deep learning; NN-based surrogate.

Introduction

Different length-scales in inertial MEMS



1. Decoupling between macro- and meso-scale ?

Allowed by small inertia of the sensor. Interaction between different scales (in case of dynamics) is **driven by the by the mass because of inertial forces.**

2. Decoupling between meso- and micro-scale?

Not possible! For a reliability assessment, particular focus has to be given to the slender parts in these inertial systems (most prone to fail), whose response strongly **depends in the underlaying microstructure.**

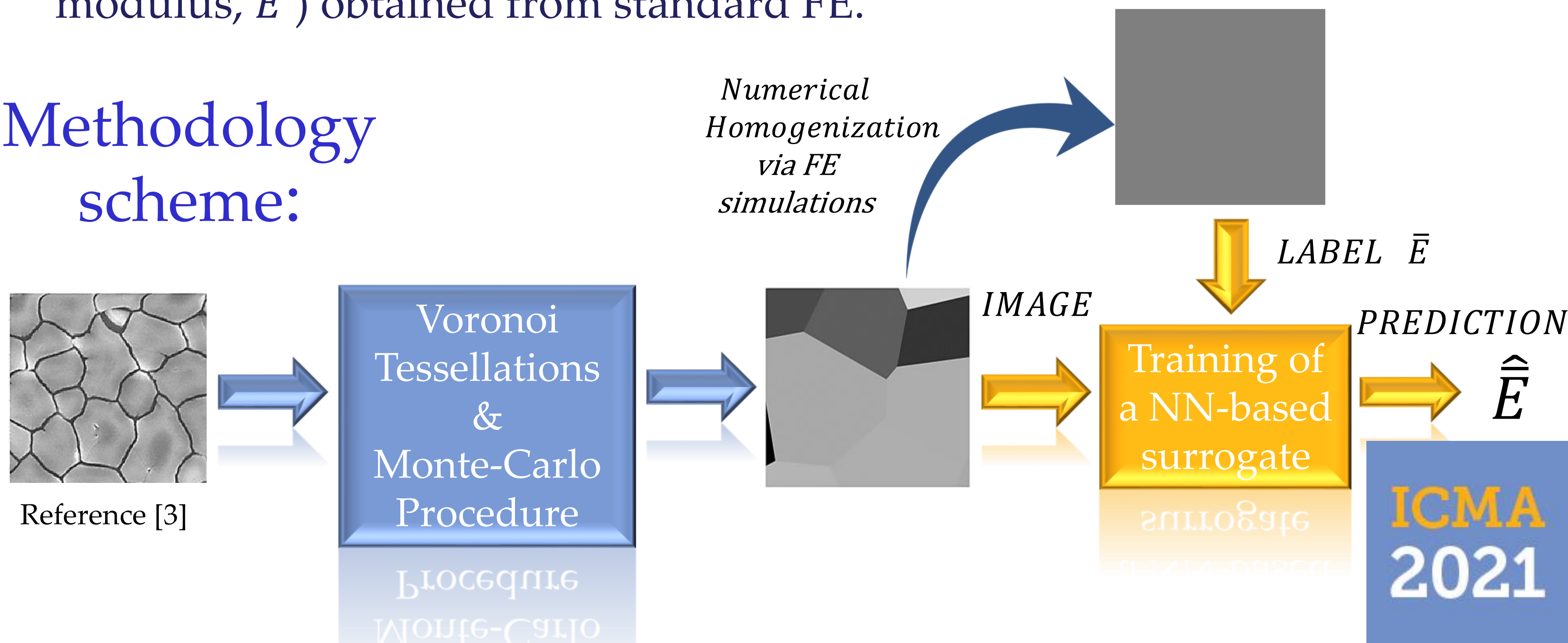
3. **Our goal is** to predict the **maximum scattering** that can be expected in the results (in the context of **homogenization**) exclusively due to intrinsic features present in the polycrystalline material constituting these devices.

ICMA 2021

Objectives

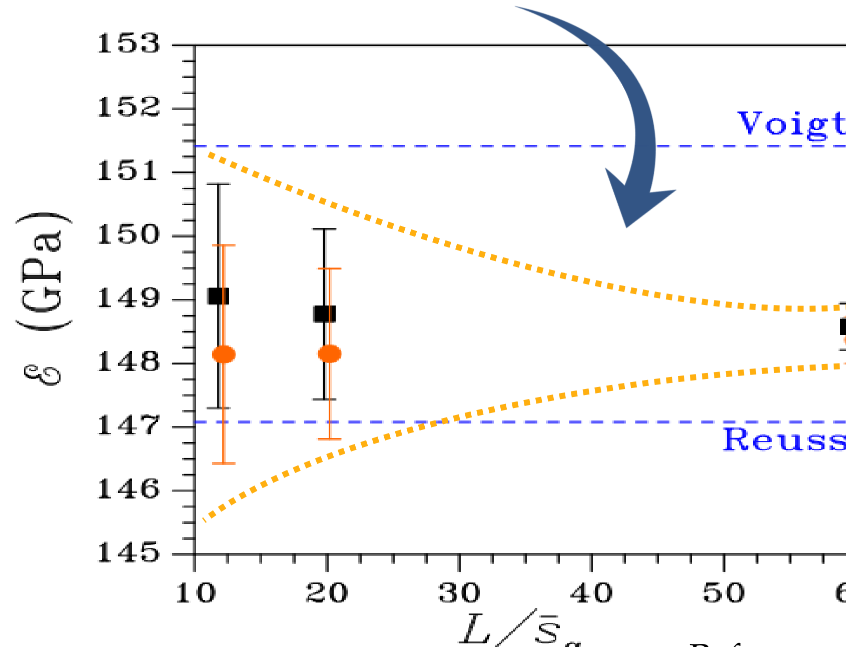
1. Propose an **alternative approach** to standard homogenization techniques.
2. Generate **microstructure-property mappings** to characterize the mechanical reliability of inertial MEMS whose movable structures are made of polysilicon films.
3. Train and test a **NN-based surrogate** that combines the sequential use of a CNN and a MLP. **Input** : 2D statistically representative images. As **labels**: theoretical values of the homogenized property (in-plane *apparent* Young's modulus, $\bar{E}$) obtained from standard FE.

Methodology scheme:



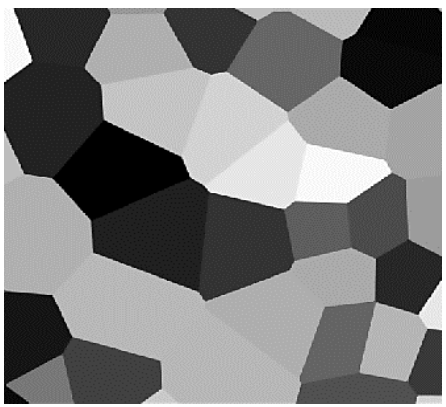
Novelty of the work

Size effects: The lower the ratio $\boldsymbol{L/\overline{s_g}}$ the larger the scattering of the overall properties around the mean



Reference [2]

Once the model has been trained, the possibility to feed images representative of different length-scales allows for a fast multiscale exploration and characterization of the **size effects** !



$$\frac{L}{\overline{s_g}} = \frac{5\mu\text{m}}{0.5\mu\text{m}} = \mathbf{10}$$

$$\frac{L}{\overline{s_g}} = \frac{3\mu\text{m}}{0.5\mu\text{m}} = \mathbf{6}$$

$$\frac{L}{\overline{s_g}} = \frac{2\mu\text{m}}{0.5\mu\text{m}} = \mathbf{4}$$

→ Trained NN-based surrogate → $\hat{\hat{E}}$

ICMA 2021

Methodology: Polysilicon Film Morphology

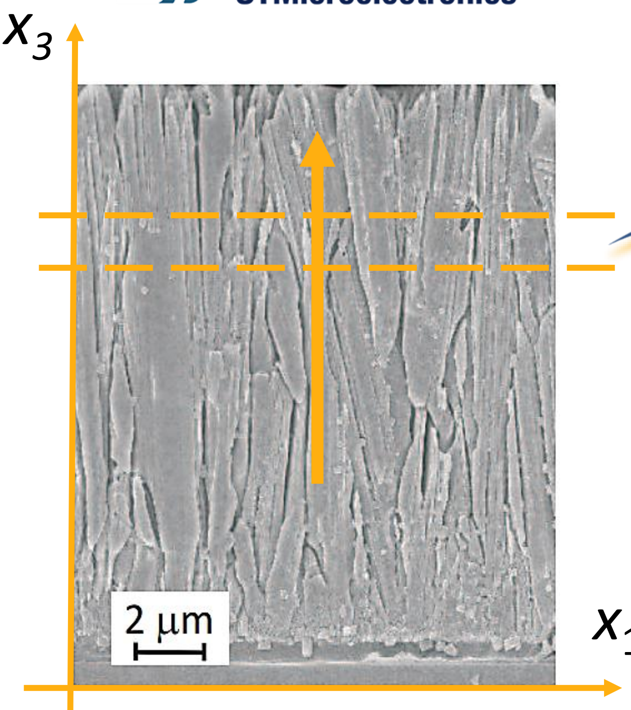
ThELMA
Surface Micromachining

STMicroelectronics

- ✓ Suspended structures
- ✓ Compliant components

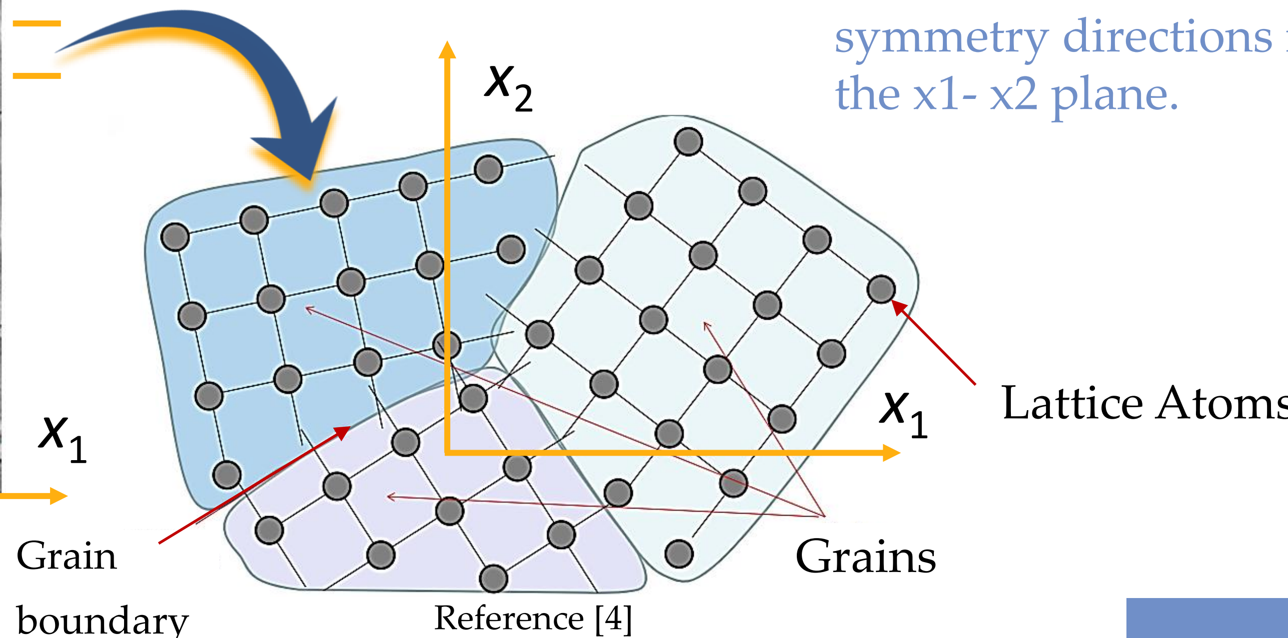
Summarizing:

1. One axis of elastic symmetry aligned with epitaxial growth direction i.e. x3.
2. Random orientation of other two elastic symmetry directions in the x1- x2 plane.



Reference [2]

Epitaxial growth
(preferential direction)



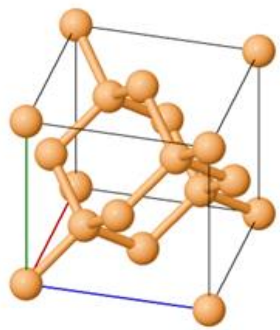
Reference [4]

Taking a slice of the film
(morphology perpendicular to growth direction)

ICMA 2021

Methodology: Monocrystalline Silicon

1. Single-crystalline silicon



Reference [5]

Diamond cubic lattice (Anisotropic elasticity)

Stiffness coefficient matrix **depends on the crystal orientation**

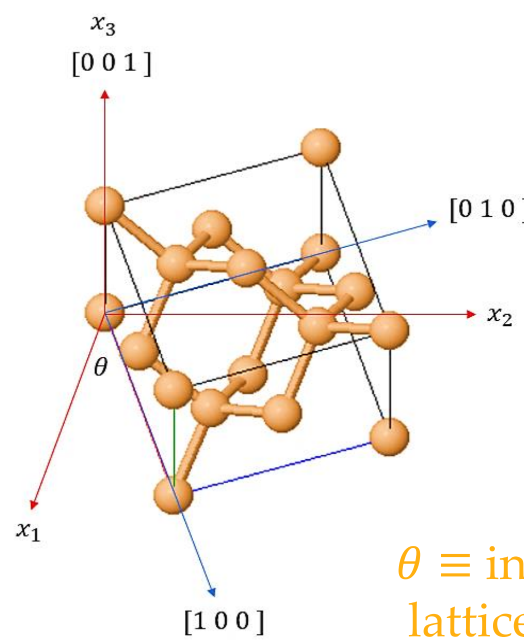
2. Stiffness Matrix of Silicon

$$\begin{pmatrix} 165.64 & 63.94 & 63.94 & 0 & 0 & 0 \\ 63.94 & 165.64 & 63.94 & 0 & 0 & 0 \\ 63.94 & 63.94 & 165.64 & 0 & 0 & 0 \\ 0 & 0 & 0 & 79.51 & 0 & 0 \\ 0 & 0 & 0 & 0 & 79.51 & 0 \\ 0 & 0 & 0 & 0 & 0 & 79.51 \end{pmatrix}$$

$< 1\ 0\ 0 >$ aligned with $(x_1\ \ x_2\ \ x_3)$
Reported in [GPa]

3. Account for lattice orientations

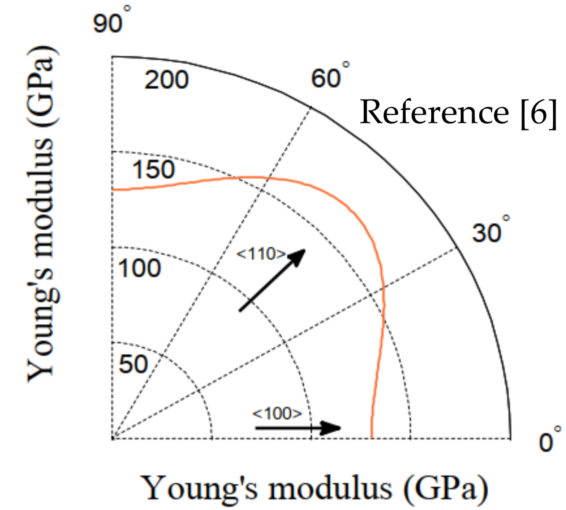
By introducing the appropriate transformation of the stiffness matrix (tensor transformation law)



$\theta \equiv$ in-plane lattice orientation

4. Directional variation of the in-plane Young's modulus

In $< 100 >$ $E = 130$ GPa
In $< 110 >$ $E = 169$ GPa



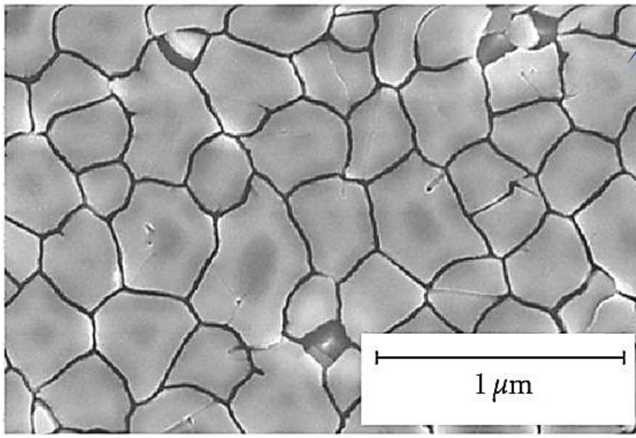
Elastic modulus of monocrystalline Si under in-plane rotations ranging $0 \leq \theta \leq 90°$

ICMA 2021

Methodology: Input Data Generation (1)

Real Microstructure

(Poly-Si Thin-Film)

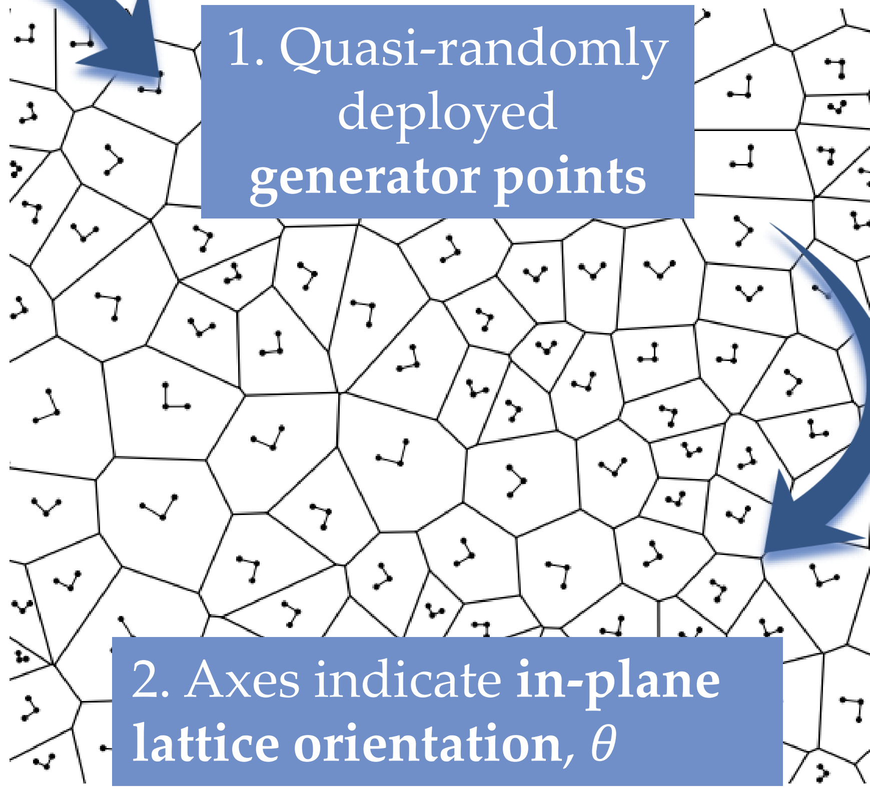


Reference [3]

Digital Microstructure

(Voronoi Tessellation)

1. Quasi-randomly deployed **generator points**



2. Axes indicate **in-plane lattice orientation**, θ

Reference [2]

Monte Carlo simulations exploited to account for the stochastic effects of:

- ✓ Topology of grain boundaries (deployment sites)
- ✓ Lattice orientations (θ)

ICMA 2021

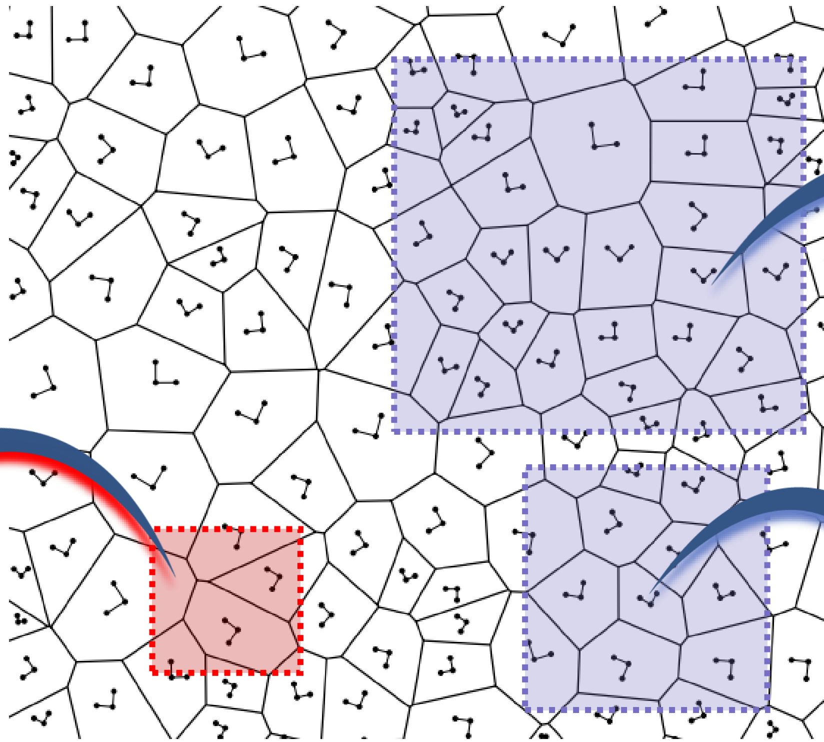
Methodology: Input Data Generation (2)

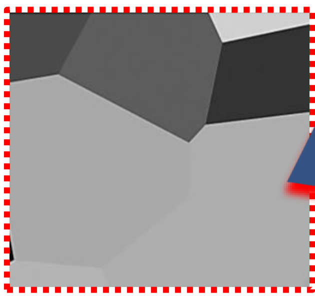
Each $L/\bar{s_g}$ ratio is linked to a different scattering. **Is our model able to catch it ?**

(*Infinite* Voronoi carpet)

$$\left.\begin{array}{l} \mathrm{L} = 2\mu\mathrm{m} \\ \bar{s_g} = 0.5\ \mu\mathrm{m} \end{array}\right\} \frac{L}{\bar{s_g}} = 4$$

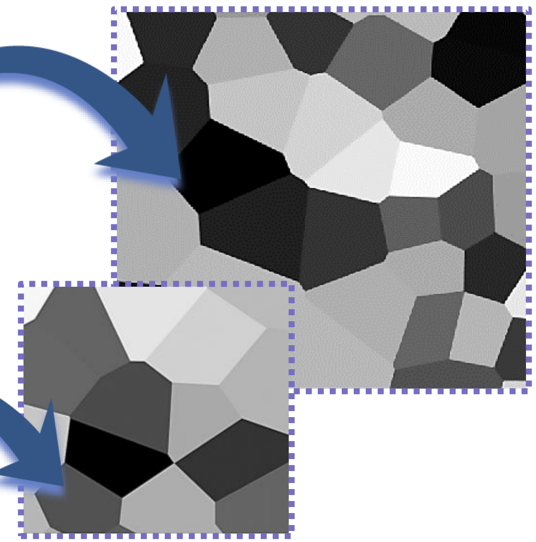
$$\left.\begin{array}{l} \mathrm{L} = 5\ \mu\mathrm{m}\ , 3\mu\mathrm{m} \\ \bar{s_g} = 0.5\ \mu\mathrm{m} \end{array}\right\} \frac{L}{\bar{s_g}} = 10, 6$$





Training & Validation Sets

Reference [2]



Two Test Sets (Assessment of the **generalization** of the model)

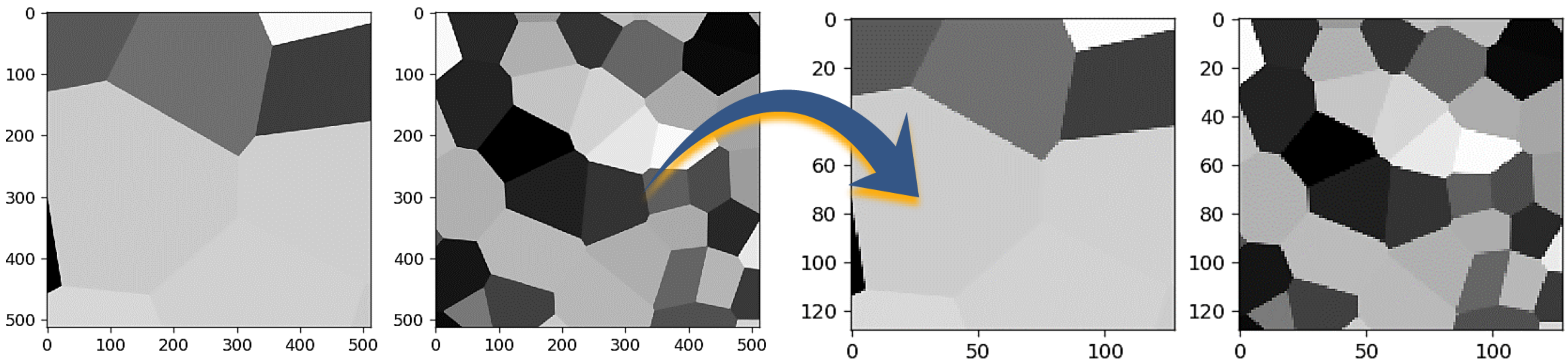
- **Pixels take values $[\mathbf{0}, \mathbf{255}]$** encoding the in-plane lattice orientation θ (related to the directional variation of the in-plane Young's modulus)
- **Ground-truth data** (labels) for SVEs come from standard FE simulations

Methodology: Input Data Pre-Processing

1. Median Filter to reduce artifacts (pixels with incorrect values)
2. Resolution Adjustment: reduction of image size (pixels) by a factor of 16

Initial resolution : 512×512 pixels.

Finally adopted resolution: 128×128 pixels.



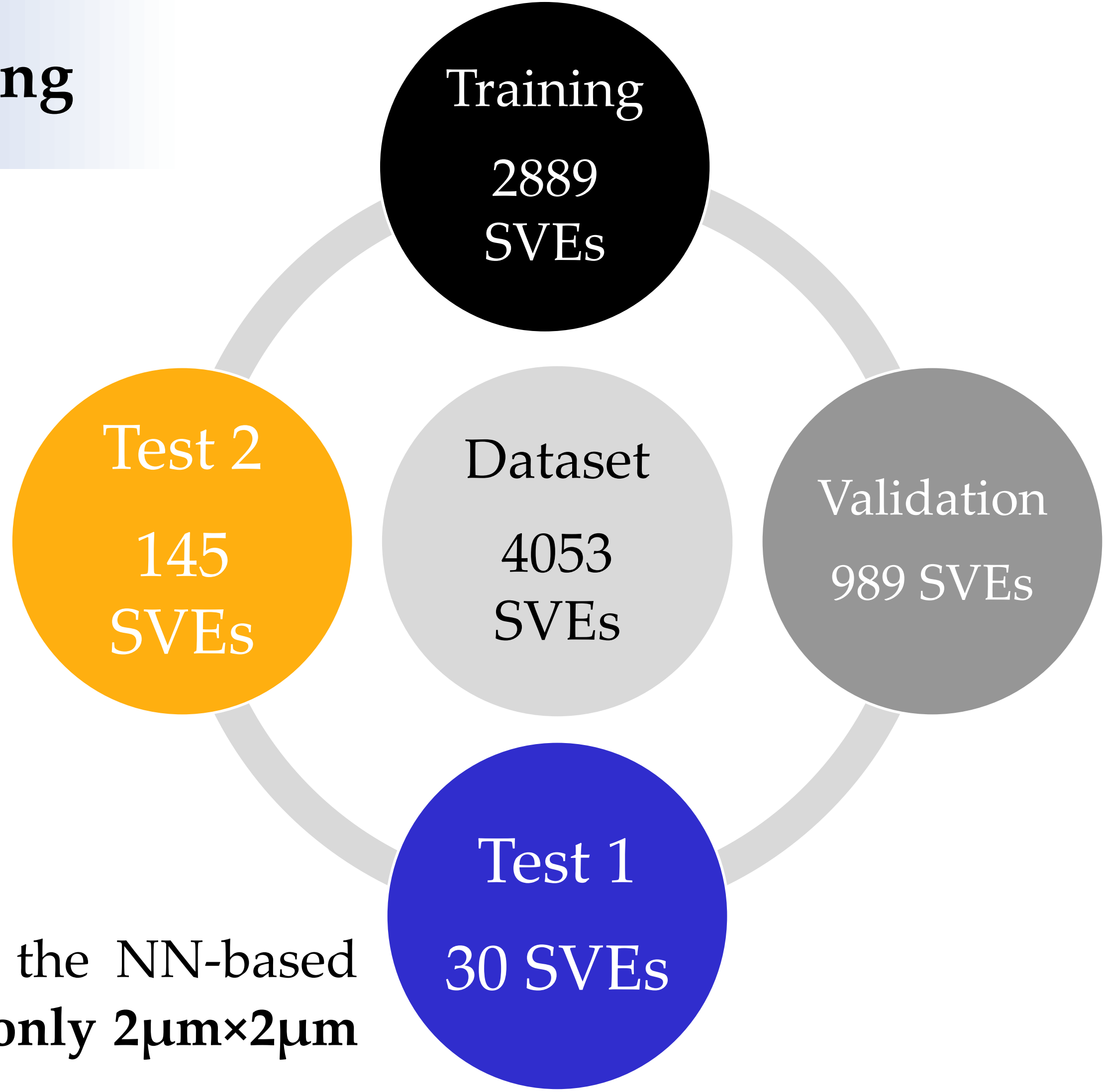
For example, for the ResNet18-based model:

Using initial resolution: ~ 72 s/epoch, Training time ~ 2.5 hours, Max. BS=10
Using final resolution: ~ 4 s/epoch, Training time ~ 8 min, Max. BS=300

Speed up the training without sacrificing model accuracy!
(test error was checked)

ICMA 2021

Methodology: Data Splitting



2μm×2μm

2μm×2μm

3μm×3μm

5μm×5μm

1. Training and validation of the NN-based surrogate is performed **using only 2μm×2μm SVE samples!**

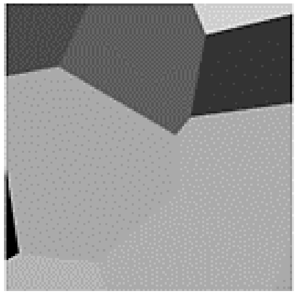
2. At a later stage, **generalization of the model** is assessed by evaluating the predictions over **3μm×3μm and 5μm×5μm SVE samples!**

Methodology: Model Implementation (1)

- Neural Network Architecture: Symbolic description of the models

Input

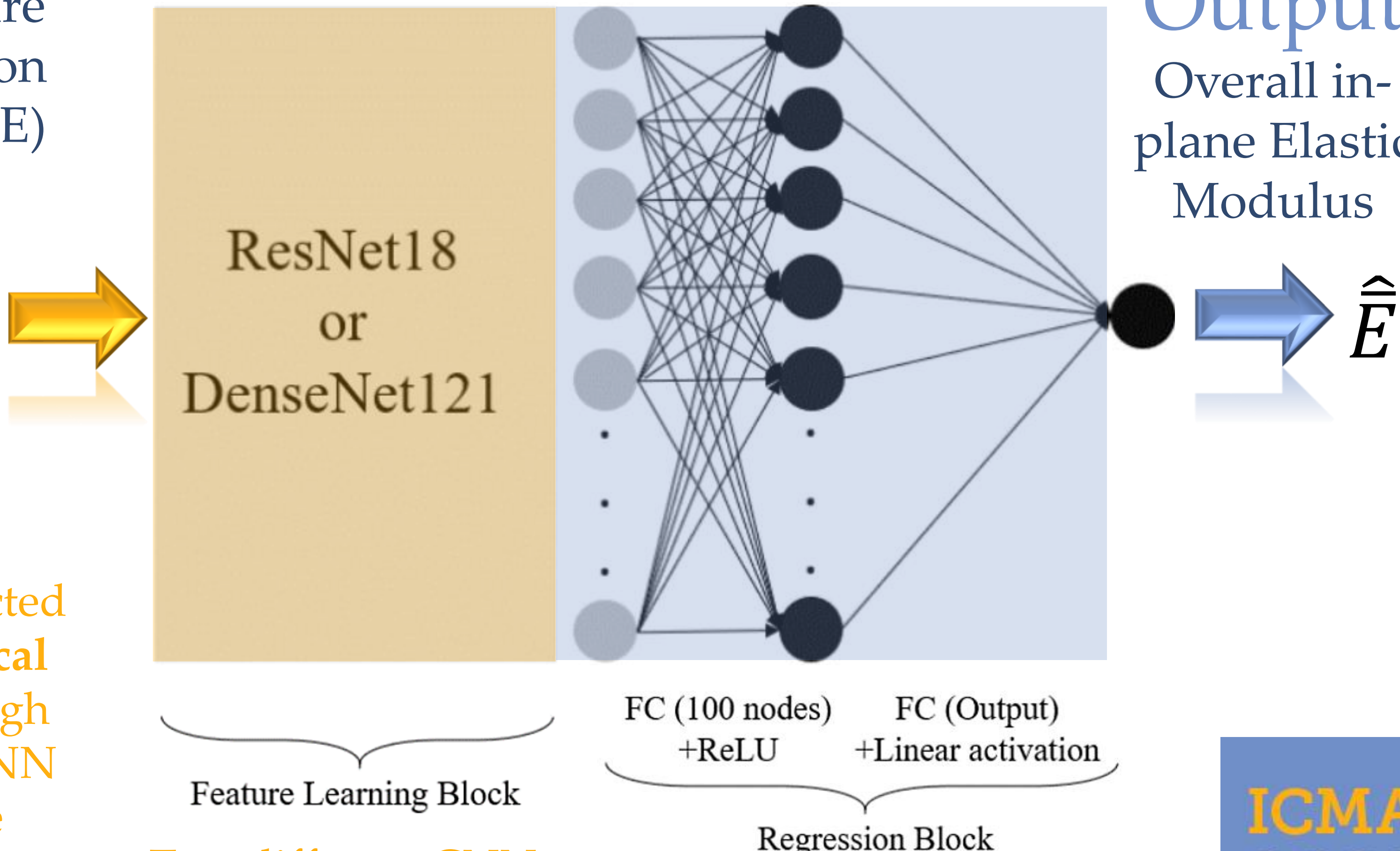
Microstructure Representation (Image of SVE)



High-level features are then employed as input to a MLP

Output

Overall in-plane Elastic Modulus



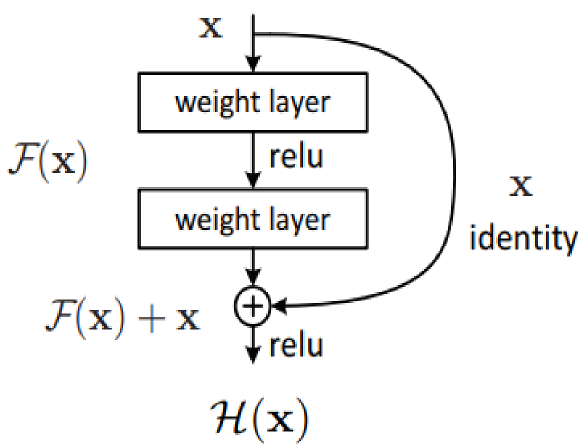
Features extracted in a **hierarchical manner** through the use of a CNN architecture

Two different CNN architectures are compared

Methodology: Model Implementation (2)

- CNN Architecture: Residual Networks

Residual learning framework based in **skip-connections**



Reference [7]

Family of ResNet Architectures

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112		7×7, 64, stride 2			
conv2_x	56×56		3×3 max pool, stride 2			
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1		average pool, 1000-d fc, softmax			
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Reference [7]

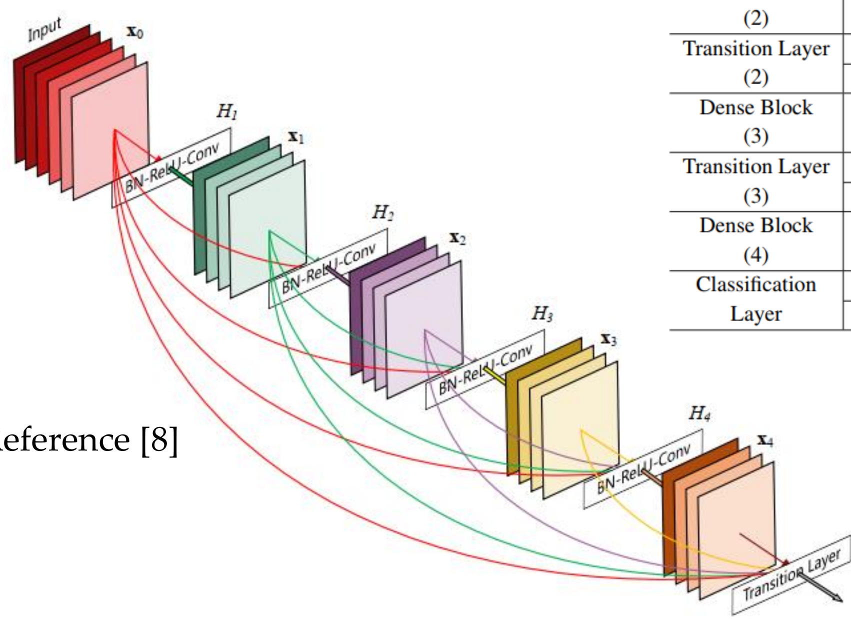
$\mathcal{H}(\mathbf{x})$ Desired underlying mapping to be fit

$\mathcal{F}(\mathbf{x}) := \mathcal{H}(\mathbf{x}) - \mathbf{x}$ Explicit fitting of a **residual mapping**

ICMA 2021

Methodology: Model Implementation (3)

- CNN Architecture: Densely Connected Convolutional Networks



Reference [8]

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Reference [8]

Family of DenseNet Architectures

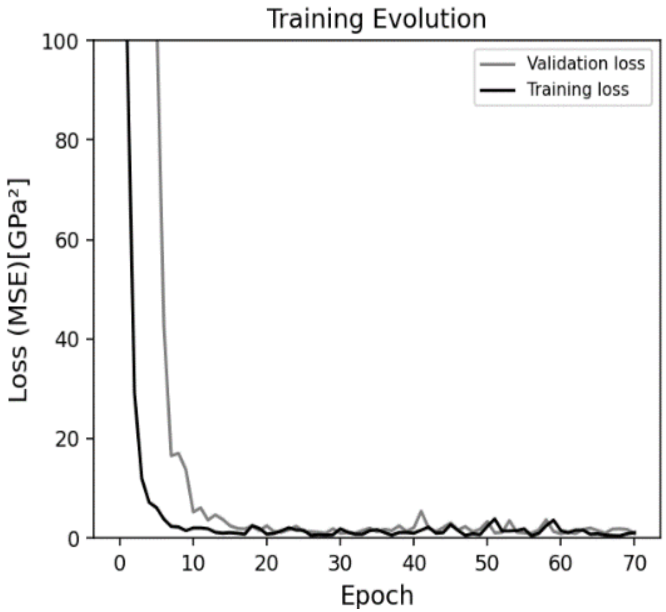
Direct connections from any layer to all subsequent layers!

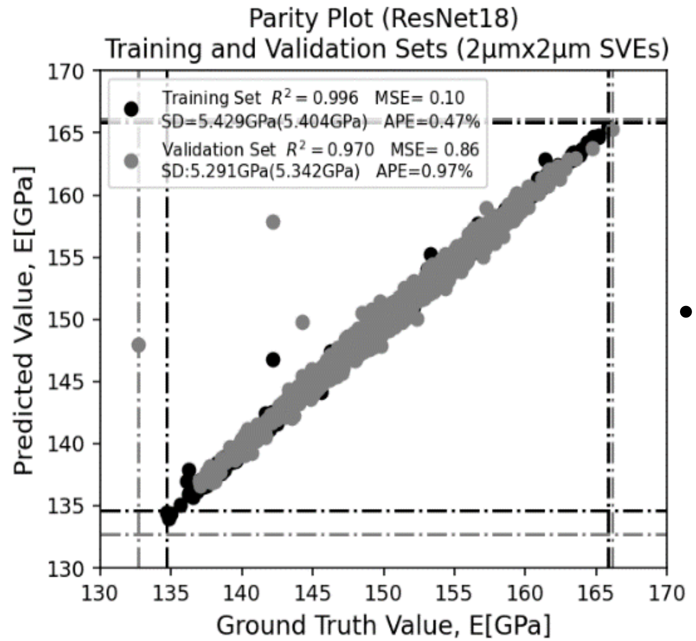
$$\mathbf{x}_\ell = H_\ell([\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}])$$

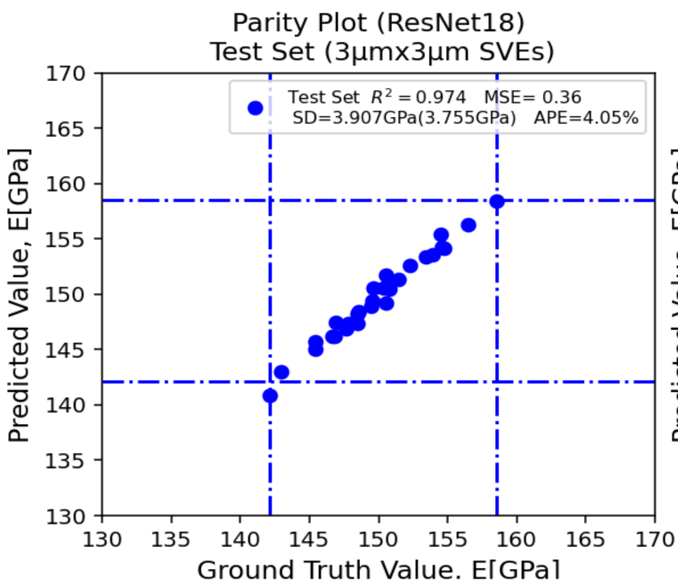
Each layer **has access to all the preceding feature-maps**, encouraging **feature reuse** (compactness, redundancy)

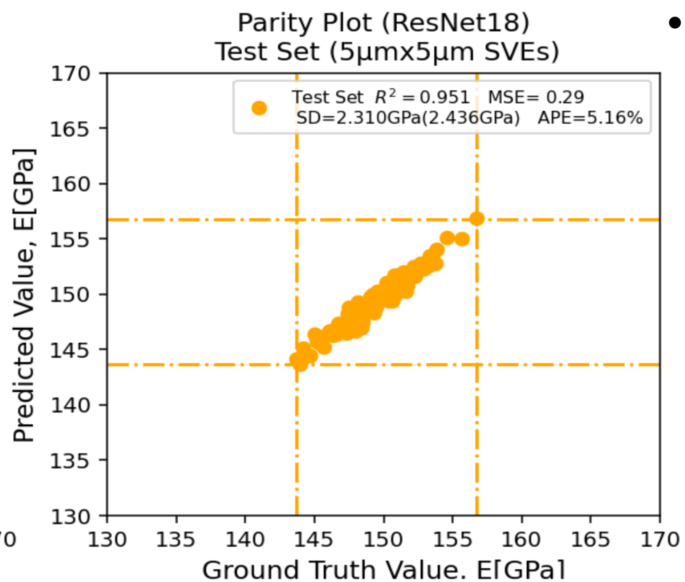
ICMA 2021

Results: ResNet-based regression model





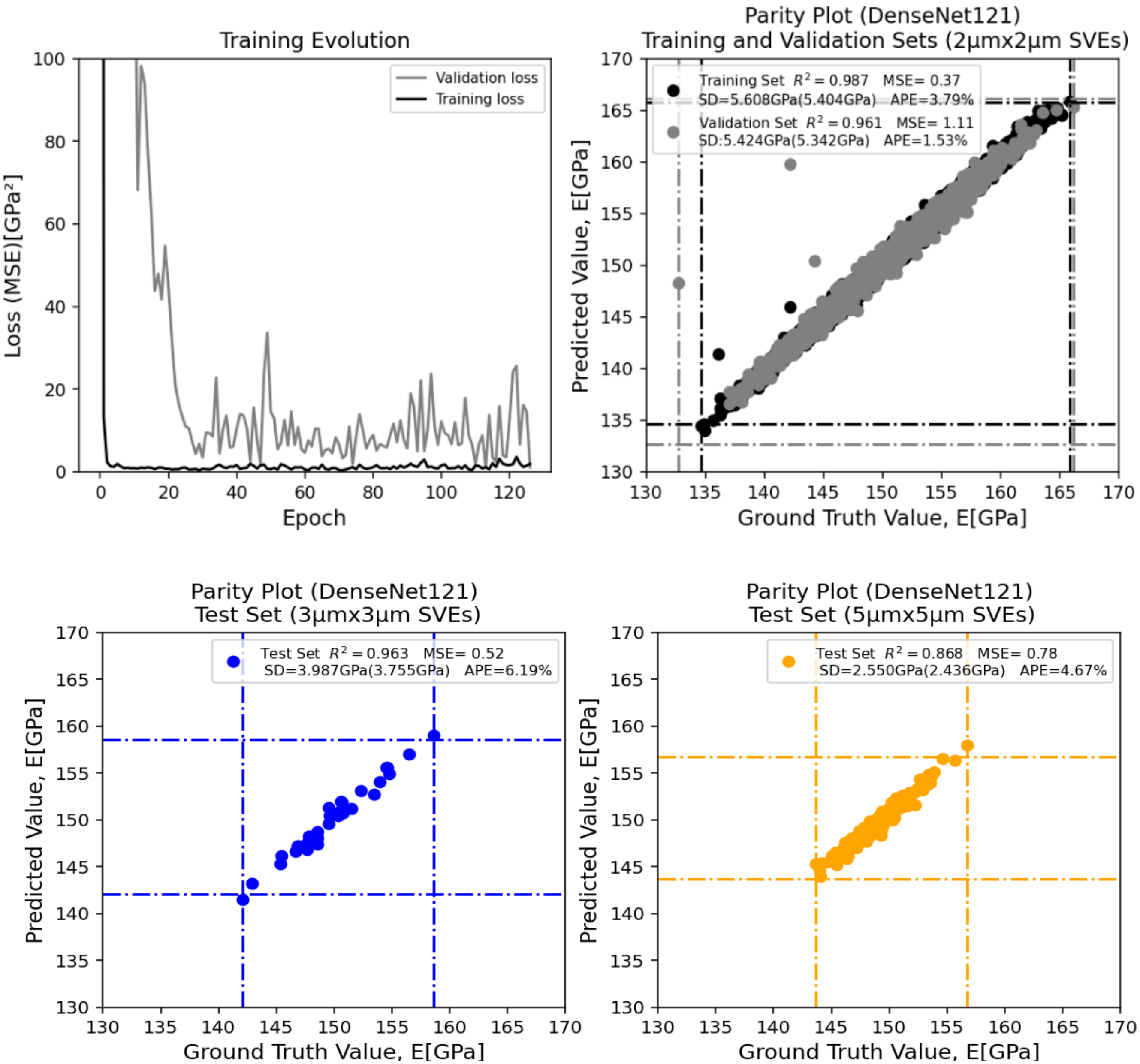




- For an ideally trained model, **data should map the identity function**, with all the dots aligned along the 45° diagonal.
- Network has learnt to **emulate intrinsic features** of the polysilicon microstructure.
- **Consistent microstructure-property mappings**: predictions fall within the theoretical limits given by the ground-truth data.

ICMA 2021

Results: DenseNet-based regression model



- Again **consistent microstructure-property mappings** are obtained.

- **Although linked to a lower total number of parameters**, this model has not displayed significant performance improvements, when compared to the ResNet18-based.

ICMA 2021

Conclusions

- The models were able to reconstruct the one-to-one correspondence between microstructural arrangements of a polysilicon aggregate and its *apparent* overall Young's modulus value.
- Statistical characterization of overall Young's modulus was possible also for SVEs featuring different sizes ($L/\overline{s_g}$ ratios) with respect to the ones employed during the training.
- Although the DenseNet121-based model requires fewer parameters, the computational time was higher than the ResNet18-based model. Moreover, DenseNet121-based model allowed the use of only a fraction of the mini batch size when compared to the ResNet18-based model.
- As far as the generalization capabilities are concerned, in general terms a better performance has been observed adopting the ResNet18-based architecture.

Prospects for future works:

Improve representativeness of test sets e.g. increase their size.

Improve pixel color encoding: account for the fact that directional variation of the in-plane Young's modulus does not follow a linear relationship with the in-plane lattice orientation.

Acknowledgements:

JPQM acknowledges the financial support provided by the University of Costa Rica (UCR) for the postgraduate studies abroad.

Partial financial support from MDPI is also gratefully acknowledged.

ICMA 2021

Supplementary Materials

- S. Mariani, R. Martini, A. Ghisi, A. Corigliano and M. Beghi, "OVERALL ELASTIC PROPERTIES OF POLYSILICON FILMS: A STATISTICAL INVESTIGATION OF THE EFFECTS OF POLYCRYSTAL MORPHOLOGY," *Journal for Multiscale Computational Engineering,* vol. 9, no. 3, pp. 327-346, 2011

- Quesada-Molina, J.P., Rosafalco, L., Mariani, S.: Stochastic mechanical characterization of polysilicon MEMS: a Deep Learning approach. In: 6th International Electronic Conference on Sensors and Applications. Proceedings. **42**, 8 (2020). doi: 10.3390/ecsa-6-06574

- Quesada-Molina, J.P., Rosafalco, L., Mariani, S.: Mechanical characterization of polysilicon MEMS devices: a stochastic, Deep Learning-based approach. In: 2020 21st International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), pp. 1-8, IEEE Press, New York (2020). doi: 10.1109/EuroSimE48426.2020.9152690

- Bagherinia, M., Mariani, S.: Stochastic Effects on the Dynamics of the Resonant Structure of a Lorentz Force MEMS Magnetometer. Actuators. **8**, 36 (2019). doi: 10.3390/act8020036

- Mirzazadeh, R., Mariani, S.: Uncertainty quantification of microstructure-governed properties of polysilicon MEMS. Micromachines. 8, 248 (2017). doi: 10.3390/mi8080248

- Mirzazadeh, R., Ghisi, A., Mariani, S.: Statistical Investigation of the Mechanical and Geometrical Properties of Polysilicon Films through On-Chip Tests. Micromachines. **9**, 53 (2018). doi: 10.3390/mi9020053

- Mariani, S., Ghisi, A., Mirzazadeh, R., Eftekhar Azam, S.: On-Chip Testing: A Miniaturized Lab to Assess Sub-Micron Uncertainties in Polysilicon MEMS. Micro and Nanosystems. **10**, 84-93 (2018). doi: 10.2174/1876402911666181204122855

- Mirzazadeh, R., Eftekhar Azam, S., Mariani, S.: Mechanical Characterization of Polysilicon MEMS: A Hybrid TMCMC/POD-Kriging Approach. Sensors. **18**, 1243 (2018). doi: 10.3390/s18041243

- Ghisi, A., Mariani, S.: Effect of imperfections due to material heterogeneity on the offset of polysilicon MEMS structures. Sensors. **19**, 3256 (2019). doi: 10.3390/s19153256

- Mariani, S., Ghisi, A., Corigliano, A., Martini, R., Simoni, B.: Two-scale simulation of drop-induced failure of polysilicon MEMS sensors. Sensors. **11**, 4972-4989 (2011). doi: 10.3390/s110504972

- Ghisi, A., Mariani, S., Corigliano, A., Zerbini, S.: Physically-based reduced order modelling of a uni-axial polysilicon MEMS accelerometer. Sensors. **12**, 13985-14003 (2012). doi: 10.3390/s121013985

- Corigliano, A., Ardito, R., Comi, C., Frangi, A., Ghisi, A., Mariani, S.: Mechanics of Microsystems. John Wiley & Sons, Hoboken, NJ, USA, (2018)

- Weinberg, M.S., Kourepenis, A.: Error sources in in-plane silicon tuning-fork MEMS gyroscopes. Journal of Microelectromechanical Systems. 15, 479–491 (2006). doi: 10.1109/JMEMS.2006.876779

- De Laat, M., Pérez Garza, H., Herder, J., Ghatkesar, M.: A review on in situ stiffness adjustment methods in MEMS. Journal of Micromechanics and Microengineering. 26, 1-21 (2016). doi: 10.1088/0960-1317/26/6/063001

ICMA 2021

Cited References in this Presentation

- [1] Kazusuke Maenaka, "MEMS inertial sensors and their applications," 2008 5th International Conference on Networked Sensing Systems, IKanazawa, Japan, 2008, pp. 71-73, doi: 10.1109/INSS.2008.4610859.
- [2] S. Mariani, R. Martini, A. Ghisi, A. Corigliano and M. Beghi, "OVERALL ELASTIC PROPERTIES OF POLYSILICON FILMS: A STATISTICAL INVESTIGATION OF THE EFFECTS OF POLYCRYSTAL MORPHOLOGY," *Journal for Multiscale Computational Engineering,* vol. 9, no. 3, pp. 327-346, 2011.
- [3] K. Kitahara, T. Ishii, J. Suzuki, T. Bessyo and N. Watanabe, "Characterization of Defects and Stress in Polycrystalline Silicon Thin Films on Glass Substrates by Raman Microscopy," International Journal of Spectroscopy, vol. 2011, pp. 1-14, 2011.
- [4] T. J. Kaiser, "Lecture 06 Solar Cell Materials & Structures," Montana State University, [Online]. Available: https://slideplayer.com/slide/6348822/.
- [5] Graz University of Technology , "Silicon," [Online]. Available: http://lampx.tugraz.at/~hadley/memm/materials/silicon/silicon.php.
- [6] R. Mirzazadeh, “Micromechanical characterization of polysilicon films: on-chip testing, multi-uncertainty quantification and Bayesian inverse modelling”, (Doctoral Dissertation). Politecnico di Milano, Milano, 2017.
- [7] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition“, 2015.
- [8] G. Huang , Z. Liu , L. van der Maaten, & K. Q. Weinberger. “Densely Connected Convolutional Networks”, 2018.