FACTA UNIVERSITATIS
Series: **Mechanical Engineering**
https://doi.org/10.22190/FUME240215019N

Original scientific paper

MACHINE LEARNING OF STRUCTURE – PROPERTY RELATIONSHIPS: AN APPLICATION TO HEAT GENERATION DURING PLASTIC DEFORMATION

Filip Nikolić[1,2], Marko Čanađija[2]

[1]CAE Department, Elaphe Propulsion Technologies Ltd, Ljubljana, Slovenia
[2]University of Rijeka, Faculty of Engineering, Rijeka, Croatia

Abstract. *In the present work, the heat generation during the plastic deformation of a multiphase material is studied using machine learning (ML) methods. The aim was to predict the temperature increase from the structure-property relationships (SPR) of a microstructure considering various Taylor–Quinney coefficients (TQCs), with the aim of achieving precision and computational efficiency suitable for industry. Using automatic microstructure generation to create datasets and finite element analysis (FEA) to obtain temperature increase – strain curves, the dataset facilitated the training of an ML model. A 3D convolutional neural network (CNN) was developed using the microstructural configuration and TQC value as input and the temperature increase – strain curve as output. The model demonstrated high prediction accuracy. The results indicated that the hard phase fraction significantly impacts the temperature increase, much more than the TQC values. This underlines the potential of the model for a better understanding of material behavior during deformation and its industrial applicability.*

Key words: *Deep learning, Taylor – Quinney coefficient, Heat generation, Structure – property relationship, Finite element analysis*

1. INTRODUCTION

It is well-known that the deformation of most materials can be divided into elastic and plastic deformation. Both can be either rate dependent or rate independent. While elastic deformation is a reversible process, plastic deformation is irreversible. During plastic deformation, most of the energy is dissipated into heat while during elastic deformation the heat generation is often neglected. For a general overview of the phenomenon, the interested reader is referred to [1] and the references therein. The issue is further complicated by the fact that the response of the material involves the temperature

Received: February 15, 2024 / Accepted May 02, 2024
Corresponding author: Marko Čanađija
University of Rijeka, Faculty of Engineering, Vukovarska 58, Rijeka, 51000, Croatia.
E-mail: marko.canadija@riteh.uniri.hr

dependence of the properties [2–4]. The fraction of plastic energy that is converted to heat is described by the Taylor–Quinney coefficient (TQC) and can depend on the type of material, strain rate, amount of plastic strain and many other factors, see [5-7]. It is often assumed to be around 90 % [5] and usually increases with deformation, see for example [2, 5, 7, 8]. Moreover, in some special cases this coefficient can be greater than one due to microstructural transformations. At this point, it should only be briefly noted that the microstructural aspects are very important, but the complexity of the phenomenon hinders the possibility of fully addressing all the details using classical techniques. Standard multiscale methods can provide some solutions to these issues, but on the other hand they are quite slow [9].

In the last few years, machine learning (ML) has become a very attractive topic in materials science and many different ML algorithms are widely used. Shallow algorithms such as the support vector machine can be found in applications such as microstructure classification [10]. Azimi et al. [11], on the other hand, used much more complex algorithms such as convolutional neural networks (CNNs) for similar tasks. DeCost et al. [12] applied a CNN model based on segmentation for novel automated applications of microstructure segmentation. Applications such as defect detection and determination of distances between secondary dendrite arms using CNNs can also be found in [13-15]. Optical microstructure images can be related to mechanical properties (yield strength, elongation and ultimate tensile strength) (see [16]). Tagimalek et al. [17] used relevance and support vector machine to determine the mechanical properties based on the input process parameters of the friction stir welding process. Similarly, Milićević et al. [18] predicted the mechanical properties of the 3D printing material PLA using neutral networks. Zarezadeh et al. [19] investigated the corrosion mechanical and electromechanical behavior of Ni coatings using a neural network and an adaptive neuro-fuzzy inference system. Li et al. [20] proposed a transfer learning approach for microstructure reconstruction and structure–property relationships (SPR).

SPR is a highly attractive topic in engineering. It can be determined using many methods, but it is most commonly determined using finite element analysis (FEA). The most significant issue with FEA is the lack of computational efficiency when it comes to simulations over two or more scales. This makes it very difficult for the industry to adopt. For this type of problem, ML methods could be of great advantage. Jung et al. [21] for example, have shown that a CNN is tremendously fast in predicting mechanical properties in the SPR domain compared to FEA simulations. Yang et al. [22] employed a CNN to predict the elastic mechanical properties of a high–contrast composite. The input for the CNN was the arrangement of the microstructure and it outperformed all other methods. Cecen et al. [23] also developed a CNN to predict mechanical properties based on the microstructure arrangement. They also used principal component analysis (PCA) for dimension reduction. Dimension reduction techniques could reduce the size of the model input, which in turn could reduce the computational cost of training. However, one should be cautious when using such methods as they may compromise the overall prediction accuracy. Wang et al. have presented an ML approach for the prediction of material properties in polymer nanocomposites. Latypov and Kalidindi [24] performed a variety of FEA simulations of digitally generated microstructures to create the dataset for training the ML model. Their model successfully predicted the yield strength and strain rate distribution in multiphase materials. To simplify the training and the ML model, they used PCA techniques for dimensionality reduction. Jung et al. [21] also created an SPR ML model

that can predict the elongation, ultimate tensile strength and strain localization index based on the arrangement of the microstructure. In addition, they used PCA and multidimensional scaling to simplify the ML model. As for the coupled thermoplasticity problems, King et al. [25] developed a physically informed machine learning model that can effectively capture the temperature increase during Shear Assisted Processing and Extrusion ShAPE. However, the work does not explicitly model the temperature increases based on the detailed microstructural characteristics or other direct material property relationships. Instead, the complex interplay between process parameters and temperature dynamics is captured by the ML model. Pantalé et al. [26] used an ML model to accurately replace the analytical formulation of a Johnson–Cook behavior law in explicit finite element simulations. Temperature is included in the formulation, but its influence manifests itself by affecting the flow law. Several different network architectures were tested. In the context of predicting stresses and temperature in the thermoelasticity of rubber, the neural network based on the strain invariants shows better performance than the network based on the stress-strain model [27]. However, plastic deformation is not relevant in such cases and was not taken into account.

From the above literature review, it can be concluded that although there are some similar research works on SPR, none of them deals with the ML prediction of TQC and heat generated during plastic deformation. To this end, the present study aims for the first time to directly predict the temperature increase curve due to plastic deformation by SPR using ML methods. The proposed methodology is based on the particular microstructural configuration of the constituents involved and their thermomechanical properties.

2. METHODS

2.1 Dataset Generation

In order to establish a relationship between the microstructure of the material and the evaluated properties using ML, a large number of different microstructures should be obtained. In this study, representative volume elements (RVEs) of equiaxed 3D two-phase microstructures consisting of 26x26x26 voxels were generated utilizing the open-source software Dream.3D [28]. The microstructure generation in the software relied on a Gaussian random distribution, and only the volume fractions of the phases were varied. Specifically, the volume fractions of the second (hard) phase ranged from 0% to 15%, while the first (soft) phase ranged from 85% to 100%. To maintain simplicity, all other parameters were kept constant throughout the generation process. The resulting microstructures had hard phase particle sizes ranging from 2 mm to 13 mm, with varying quantities of hard phase particles ranging from 1 to 15. These particles were randomly distributed in the microstructure. It is noteworthy that the generation parameters employed here are consistent with those used in a previous study [29]. For a more comprehensive understanding of the parameters used, interested readers are encouraged to refer to this study. It is emphasized that the present investigation of the mechanical behavior of the material is based on a more realistic 3D scale than the usually considered 2D scale [30–36]. Nevertheless, it is important to emphasize that it was not the intention to explicitly replicate real materials, although the methodology is applicable to these as well. Instead, the aim was to investigate the impact of different particle sizes and microstructural

configurations on the mechanical properties and behavior of hypothetical composites under different conditions. This approach allows a wide spectrum of particle sizes and microstructural configurations to be explored, some of which could potentially be larger than those commonly found in real composites. Finally, 3474 different microstructure arrangements were generated. It should be noted that it is important to obtain a large number of different microstructural configurations since heat generation also depends on the configuration.

To determine the plastic behavior of each RVE, an FEA is performed using Abaqus version 2021. For the material properties, Youngs moduli of 50 and 500 GPa were used for the first and second phase, respectively, while a Poisson's ratio of 0.33 was chosen for each phase. The true (Cauchy) stress - true plastic strain curves of each phase are shown in Fig. 1.

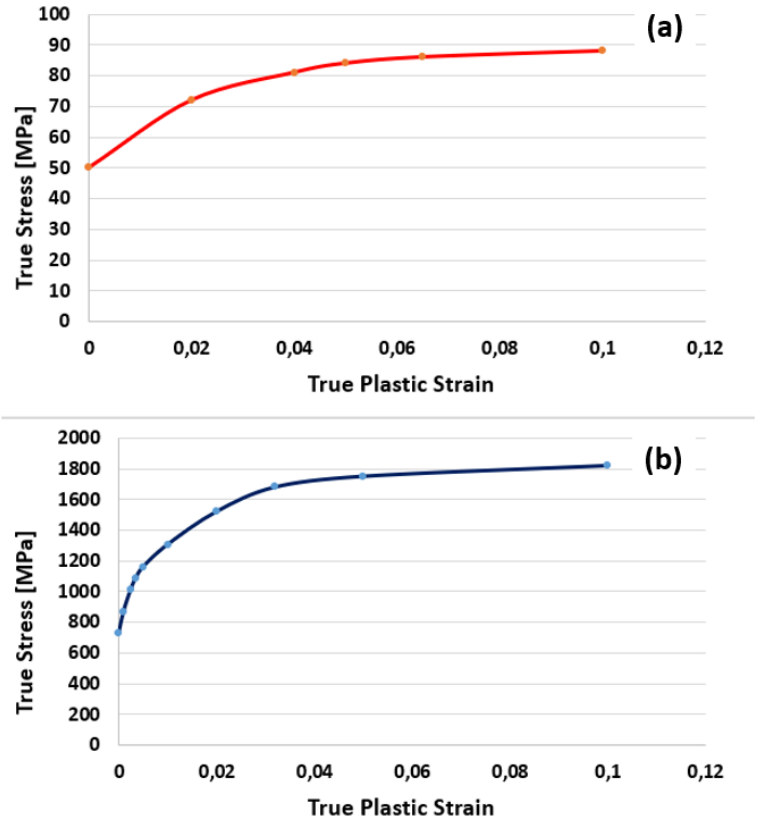

Fig. 1 True stress – true plastic strain diagrams of: (a) soft phase and (b) hard phase of the microstructure

These curves resemble typical stress-strain curves represented as a series of stress-strain points as determined in a tensile test. No attempt was undertaken to establish a specific equation relating two quantities. The mapping of the properties of each phase to a finite

element mesh, the size and element types of the mesh, and the simulation procedure matches our previous study [30]. Periodic boundary conditions were enforced, while the load was specified as an enforced displacement, resulting in a strain magnitude of 0.57%. Thus, the numerical procedure is equivalent to the macroscopic uniaxial tensile tests. The only difference was that in the present study the influence of temperature on the mechanical properties was neglected for simplicity. The entire dataset took approximately 530 hours to generate using an Intel®CoreT M I7 – 4970 central processing unit (CPU) and a single processing unit. The simulation took between 7 and 12 minutes and on average about 9 minutes, depending on the microstructure configuration. The strain magnitude was kept relatively small due to the computational effort. Very high strain magnitudes require a significant number of simulation steps due to convergence issues and drastically increase the simulation time. Therefore, a lower strain magnitude was chosen so that the simulations could be performed in a realistic time frame. The output of the analysis was the plastic strain energy, which was given in 20 evenly spaced intervals. Thus, the arrangement of the microstructures was the only variable parameter for the FEA simulation.

The arrangement of the phases with three different hard phase fractions is shown in Fig. 2. Note that the hard phase fraction varied from 0% to 0.15% in the entire dataset. The hard phase fraction was rather low as the intention was to show the effect of strengthening a weaker material with a stronger material. For each RVE in Fig. 2, the resulting true von Mises stress, the total displacement and the true equivalent plastic strain at the last simulation increment are shown for the corresponding RVEs (Figs. 3, 4, 5, 6). Note that the displacement in each figure is scaled by 25. Some elements of the RVE are intentionally hidden to visualize the material behavior within the RVE. It should also be noted that the von Mises stress increases as the fraction of the hard phase increases, especially within the elements to which material properties of the hard phase are assigned. For the soft phase, only the non-uniformity of the stress field is higher, while the general level hardly changes, Fig. 3. For the highest fraction of the hard phase, the deformation is higher, while there is hardly any difference for the two lower values, Fig. 4. Both the general level and the non-uniformity of the plastic deformation increase with increasing hard phase fraction, Fig. 5. Furthermore, it was observed that the level of plastic deformation is significantly lower within the elements to which hard phase material properties are assigned, Fig. 6.

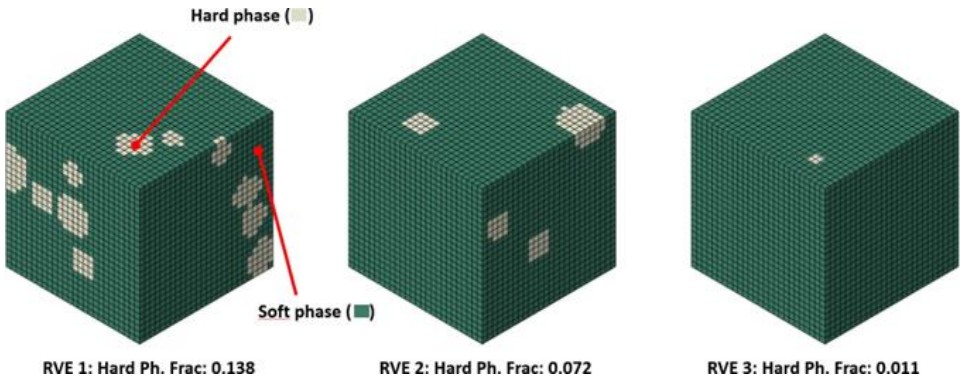

Fig. 2 Visualization of the arrangement of the phases for a different RVEs

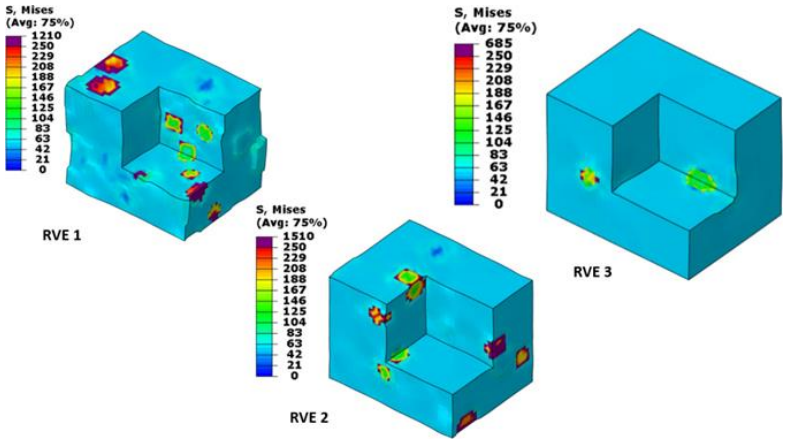

Fig. 3 Von Mises stress (MPa) for the RVEs in Fig. 2

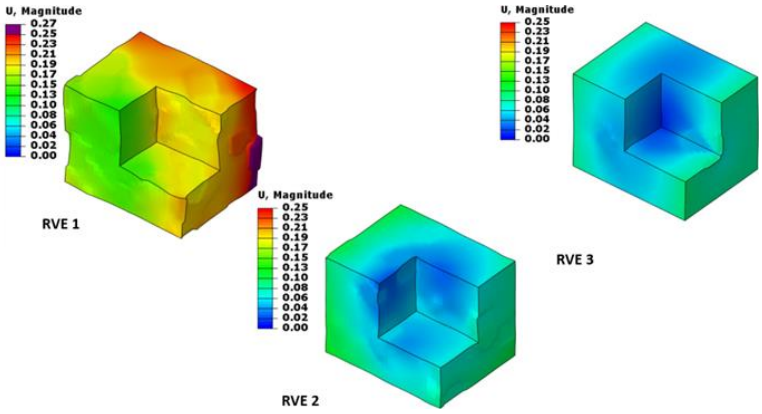

Fig. 4 Total displacement (mm) for the RVEs in Fig. 2

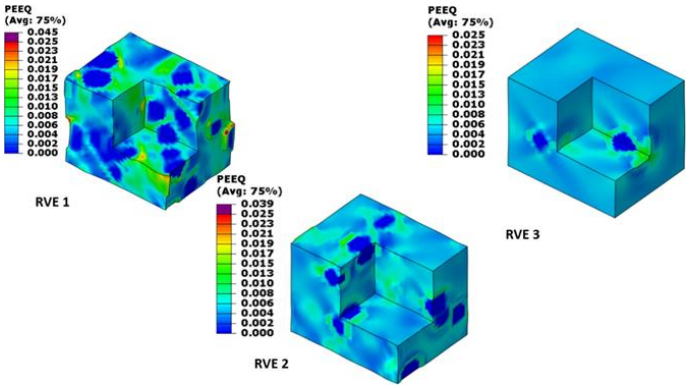

Fig. 5 Equivalent plastic strain for the RVEs in Fig. 2

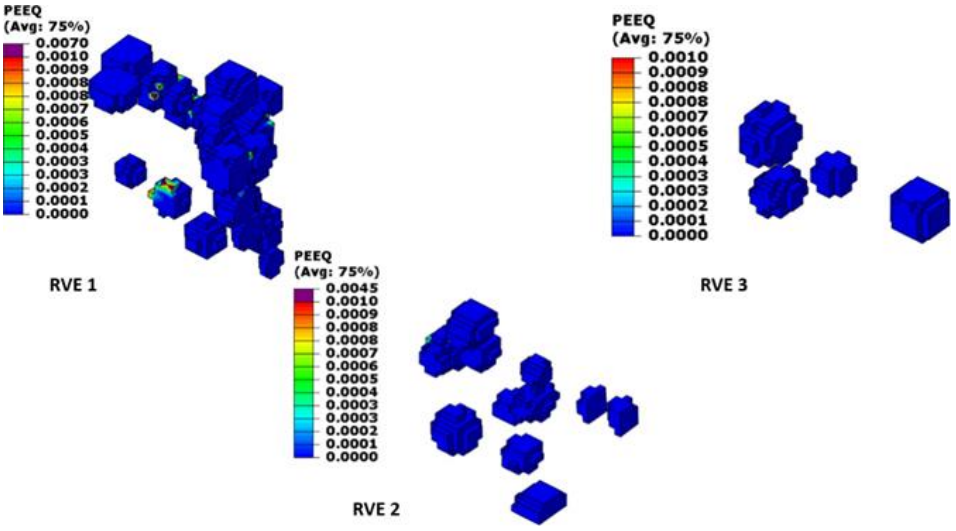

Fig. 6 Equivalent plastic strain for the RVEs in Fig. 2 (only hard phase is shown)

Knowing the amount of plastic energy dissipated during the deformation of the RVE allows the calculation of the average temperature increase in the RVE during deformation, as outlined below [1]:

$$\Delta T = \frac{Q \cdot \beta}{\rho \cdot V \cdot c_p} \tag{1}$$

where ΔT is the temperature increase in [K], Q is the total plastic energy of the RVE in [J], β is the TQC, ρ is the mass density in [kg/m^3], V is the volume of the RVE in [m^3] and c_p is the specific heat capacity in [J/kg K]. It is important to note that the above equation is only valid for adiabatic conditions, which corresponds to higher strain rates. However, the non – uniformity of the temperature field due to the non-uniform plastic deformation was not taken into account and simplification that the entire RVE heats uniformly was considered. The volume of the RVE was simply calculated based on its dimensions (26x26x26 mm) and corresponds to V = $1.7576 \cdot 10^{-5}$ m^3, while the amount of plastic deformation was obtained from the FEA simulation. The mass density and specific heat capacity were both kept constant for the dataset and are: c_p = 500 J/kg K and ρ = 2500 kg/m^3. For simplicity, the same c_p and ρ were assumed for both phases. Even though these properties were kept the same for both phases in this study, it should not be a fundamental problem to repeat a similar study with different proportions between c_p and ρ. The TQC β is designed to change based on the level of strain experienced by each RVE in the dataset:

$$\beta = \beta_1(18.86\varepsilon + 1). \tag{2}$$

The constant β_1 is randomly chosen between 0.7 and 0.8 for each RVE. For the zero-strain and the maximum tensile strain (0.57%) in the direction of the prescribed displacement, it turns out that β can fall within the range of 0.7 to 0.886. It should be noted that the Eq. (2) is not an attempt to physically describe the real evolution of the TQC, but merely serves to

promote additional complexity in the heat generation process while still keeping the TQC within reasonable limits for some widely used construction materials.

2.2 Overview of the ML Model

In general, ML algorithms can be divided into two main groups: shallow learning (e.g., smaller artificial neural networks, vector machine, decision tree, and others) and deep learning (DL) (e.g., recurrent neural networks, deep belief networks, CNN, and others) [37]. Many authors have shown that complex DL algorithms typically produce better results [11–15, 20, 38–43]. However, determining the optimal balance between computational resources and desired precision remains an open issue.

The ML model used in the present study is a complex 3D-CNN that considers input in the form of a 26x26x26 matrix. The matrix contains information about the arrangement of the hard and soft phases of the microstructure as well as the TQC value. The initial TQC value was used to create the matrix as explained in the previous subsection. The initial matrix consists of zeros (soft phase) and TQC (hard phase). Including the TQC value in this manner is a convenient way to pass information about the TQC value to the ML model and simultaneously define the distribution of the hard phase. Thus, the TQC value is directly included in the CNN. Otherwise, the CNN should be considered as a multiple – input (i.e., information about the phase arrangement and the TQC value), which is more computationally intensive. Therefore, the ML model captures any arrangement of phases in the RVE and can be used for the same purpose as the macroscopic uniaxial tensile test. Since the model was trained using data obtained from the uniaxial tensile test, application to other boundary conditions should be possible, but an initial verification on some test examples is recommended. It should be noted that only the results corresponding to ε greater than 0.00228 were used for training. This threshold corresponds to the last 12 increments, while the first 8 increments were not used. Up to ε of 0.00228, the plastic strain is very low and only occurs due to some stress concentration at the interface between the soft and hard phase. Therefore, the heat generation is also very low in contrast to the part of the curve that includes ε above 0.00228. As a consequence, the temperature increase within the first 8 points is negligible in contrast to the temperature increase within the following 12 points. Thus, excluding the first 8 points of the curve from the learning process, allows the CNN to be much simpler and computationally more efficient, while the loss of accuracy is minimal.

Additionally, several alternative CNN configurations are examined to compare the performance of the different model structures. While numerous experiments have been conducted with different CNN configurations to determine the best performing one, the focus here is limited to six distinct CNNs. The six distinct CNN models were tested on the same test dataset, and their respective performance results are subsequently presented. The performance analysis was concentrated on the maximum global prediction error, mean squared error (MSE), and root mean squared error (RMSE) metrics, Tab. 1. It should be noted that all strain increments were used to evaluate the results.

According to the Tab. 1, it can be concluded that CNN1 is the best performing model and was selected for the task at hand. A brief comparison of the performances of the different CNNs is presented in the Results and Discussion section. CNN1 was constructed using two blocks built with 3D convolutional layers. Each convolution block follows a Parametric Rectified Linear Unit (PReLU) activation, a batch normalization and a max–

pooling layer. For the first block, filter sizes of (8x8x8x32) were used, while filter sizes of (6x6x6x32) were used after the max-pooling layers. For the second block, filter sizes of (5x5x5x32) were used while after the max – pooling layer, filter sizes of (4x4x4x32) were used. The last convolution block is followed by a meta–layer, which is then followed by a fully connected layers of 512, 256, 128, 64 and 20 neurons. Each of these layers is followed by a batch normalization and the PReLU activation function. The last layer is followed by 12 logistic output neurons that output the 12 uniformly distributed temperature increase intervals. The CNN operates on the feed–forward (sequential) principle. Thus, for the microstructure arrangement and the input of the TQC value, the 12 points of the temperature increase vs. true strain curve were output. As already mentioned, dimension reduction techniques such as PCA could affect the overall efficiency of the model, so they are avoided in this study. A schematic representation of the model can be found in Fig. 7.

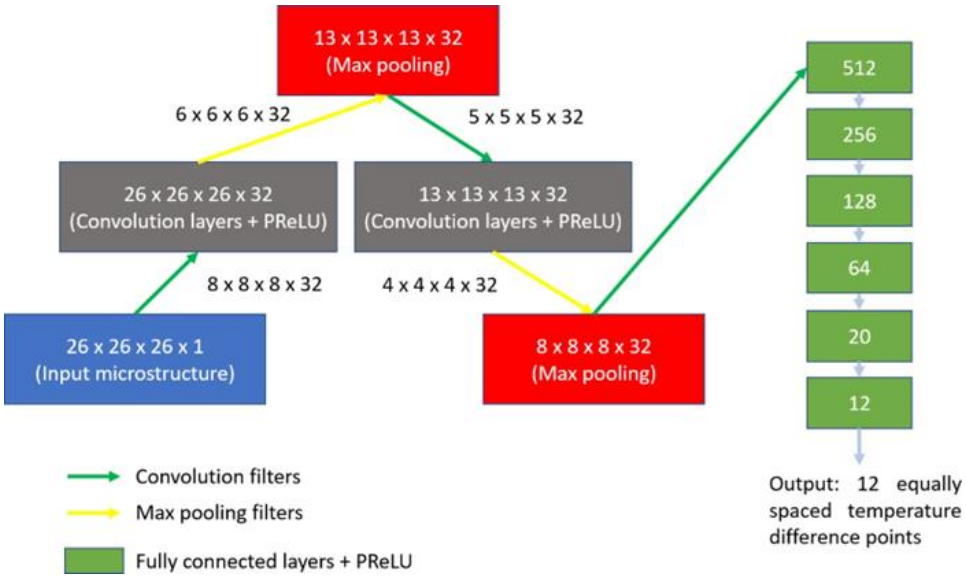

Fig. 7 A schematic representation of the machine learning model

The CNN was trained with Python version 3.9 and Tensorflow version 2.6 using the Keras library. Convergence was achieved after 300 epochs with a learning rate of 0.00025 and an Adam optimizer. The training was completed in approximately one hour using an Nvidia Quadro K2200 graphics card. The dataset for training was split into the training set, the validation set and the test set in a 70:20:10 ratio. MSE was used as the validation loss metric, and the model was validated using the curves from the FEA. It is worth emphasizing that the training procedures remained consistent for all different CNNs presented in Tab. 1, thus, only the CNN configuration itself was changed. Unlike some other studies [21, 23, 24, 43], the advantage of the present ML model was that simplifications such as dimension reduction techniques were not necessary, see Results and Discussion section. In summary, the CNN uses the microstructural configuration of the RVE and TQC as input and outputs the temperature increase–strain curves. A schematic representation of the entire learning process is shown in Fig. 8.

Table 1 CNN Models configuration and performance

Model	CNN Configuration	MSE	RMSE	Error	Comp. Time
CNN1	CL_1 (f-16, (26x26x26x32) – Cf(8x8x8x32) A-PReLU), BN, Mp_1 ((13x13x13x32) – Mpf(6x6x6x32)) CL_2 (f-16, (13x13x13x32) – Cf(5x5x5x32) A-PReLU)), BN, Mp_2 ((8x8x8x32) – Mpf(4x4x4x32)) D(512, A-PReLU), BN, D(256, A-PReLU), BN, D(128, A-PReLU), BN, D(64, A-PReLU), BN, D(20, A-PReLU), BN, Output (Opt: Adam)	1.94e-5	0.004	6 %	68 min
CNN2	CL_1 (f-16, (26x26x26x32) – Cf(5x5x5x32) A-PReLU), BN, Mp_1 ((13x13x13x32) – Mpf(2x2x2x32)) CL_2 (f-16, (13x13x13x32) – Cf(3x3x3x32) A-PReLU)), BN, Mp_2 ((8x8x8x32) – Mpf(2x2x2x32)) D(1024, A-PReLU), D(0.5), BN, D(512, A-PReLU), D(0.5), BN, D(256, A-PReLU), D(0.5), BN, D(128, A-PReLU), BN, D(64, A-PReLU), BN, D(20, A-PReLU), BN, Output, (Opt: Adam)	2.97e-5	0.005	20 %	54 min
CNN3	CL_1 (f-16, (26x26x26x32) – Cf(5x5x5x32) A-ReLU), BN, Mp_1 ((13x13x13x32) – Mpf(2x2x2x32)) CL_2 (f-16, (13x13x13x32) – Cf(3x3x3x32) A-ReLU)), BN, Mp_2 ((8x8x8x32) – Mpf(2x2x2x32)) D(1024, A-ReLU), D(0.5), BN, D(512, A-ReLU), D(0.5), BN, D(256, A-ReLU), D(0.5), BN, D(128, A-ReLU), BN, D(64, A-ReLU), BN, D(20, A-ReLU), BN, Output, (Opt: SGD)	0.01	0.094	400 %	42 min
CNN4	CL_1 (f-16, (26x26x26x32) – Cf(8x8x8x32) A-PReLU), BN, Mp_1 ((13x13x13x32) – Mpf(6x6x6x32)) CL_2 (f-16, (13x13x13x32) – Cf(5x5x5x32) A-PReLU)), BN, Mp_2 ((8x8x8x32) – Mpf(4x4x4x32)) D(512, A-PReLU), BN, D(256, A-PReLU), BN, D(128, A-PReLU), Output, (Opt: Adam)	6.52e-5	0.007	50 %	52 min
CNN5	CL_1 (f-16, (26x26x26x32) – Cf(5x5x5x32) A-ReLU), BN, Mp_1 ((13x13x13x32) – Mpf(4x4x4x32)) CL_2 (f-16, (13x13x13x32) – Cf(3x3x3x32) A-ReLU)), BN, Mp_2 ((8x8x8x32) – Mpf(3x3x3x32)) D(256, A-ReLU), BN, D(128, A-ReLU), BN, Output, (Opt: SGD)	0.03	0.17	1400 %	38 min
CNN6	CL_1 (f-16, (26x26x26x32) – Cf(5x5x5x32) A-ReLU), BN, Mp_1 ((13x13x13x32) – Mpf(4x4x4x32)) D(128, A-ReLU), BN, Output, (Opt: SGD)	0.07	0.26	2000 %	28 min
Legend:	CLx – 3D convolutional layer, f – number of filters, Cf – number of convolutional filters, A – Activation function, Mpx – Max pooling layer, Mpf – Number of max pooling filters BN – batch normalization layer, Dr – Dropout layer, D – Dense (fully connected layers), Opt - Optimizer				

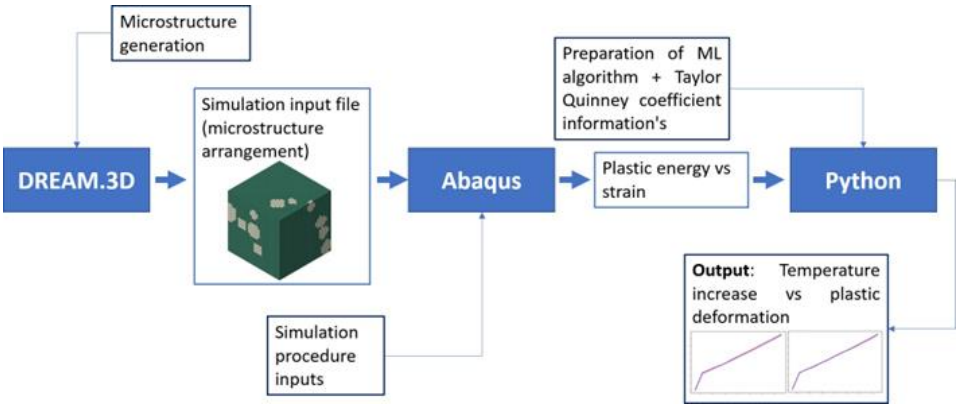

Fig. 8 Schematic representation of the complete machine learning process

3. RESULTS AND DISCUSSION

3.1 Evaluation of Different CNN Configurations

A brief comparison of the predictive capabilities of the six distinct CNNs shown in Tab 1 can be seen in Fig. 9. The CNN5 and CNN6 configurations represent an attempt to address the task at hand using an extremely basic CNN configuration. Both models show remarkable similarities and exude simplicity. Unfortunately, both performed very poorly in this task. MSE and RMSE were significantly high compared to the best performing model and the error plots showed an error margin exceeding 1000%. Therefore, it can be concluded that such simplified CNN models are not suitable for conducting the SPR task at this scale. The CNN4 model introduces a slightly more complex structure and is additionally characterized by the use of an Adam optimizer instead of SGD (stochastic gradient decent) and PReLU activation functions instead of Rectified Linear Unit (ReLU). This model demonstrated a significantly improved performance compared to CNN6 and CNN5. However, the results achieved still did not reach the standards required for the task at hand. Consequently, it can be concluded that the Adam optimizer and the combination of PReLU activation functions perform better than the combination of SGD optimizer and the ReLU activation functions for the current task.

Moreover, the more complex CNN3 model using an SGD optimizer and ReLU activation functions, performed worse compared to a simpler CNN4 model, confirming the above hypothesis. In general, it can be concluded that although the performance of the more complex CNN3 model may not be suitable for the current task, it outperforms the simpler CNN6 and CNN5 models that use the same activation functions and optimizers. The CNN2 model shares a similar configuration to the CNN1 model, although it is slightly less complex. It is also worth noting that the comparison between the CNN2 and CNN1 models leads to the conclusion that a small proportion of dropout slightly reduces the performance of the model. The CNN2 model provided commendable performance on the task at hand. However, in terms of metrics and global error, its results were inferior to those of the CNN1 model. In addition, it is also visible that increasing the amount of kernel and max-pooling filters size increased the computation time. For this reason, the CNN1 configuration is

chosen as the primary one in this study. Therefore, it can be concluded that the complexity of the model as well as the choice of activation functions and optimizers are crucial for the accurate prediction of the temperature increase-strain curves involving the SPR task. It should be emphasized again that the study did not investigate all possible combinations of activation functions, parameters and optimizers. Instead, only some of the most popular choices were presented.

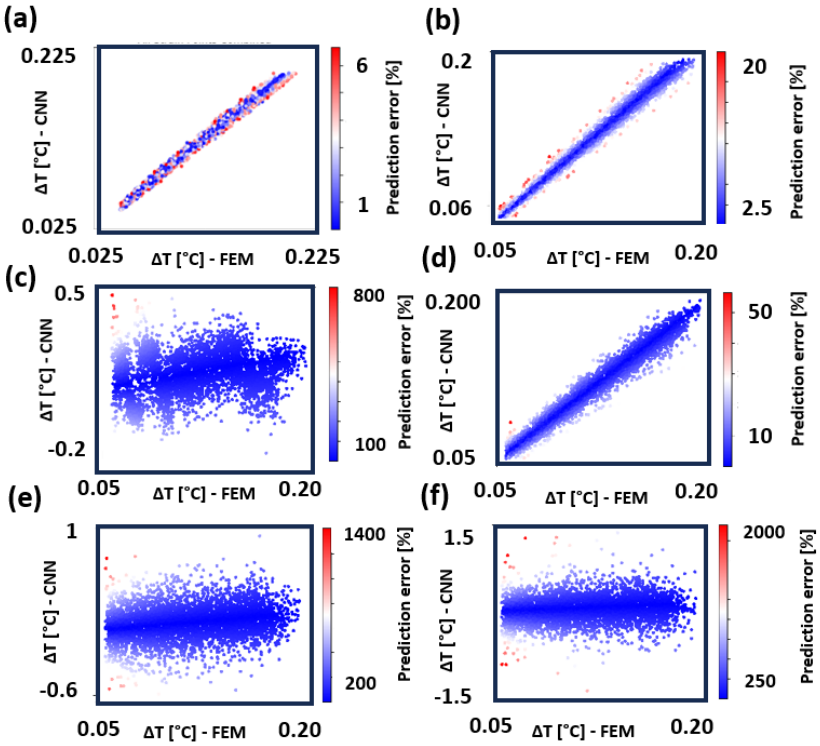

Fig. 9 Comparison between the calculated (FEA) and the predicted (CNN) value of the temperature increase for all points of the temperature increase curve for: (a) – CNN1 Model, (b) – CNN2 Model, (c) – CNN3 Model, (d) – CNN4 Model, (e) - CNN5 Model, (f) – CNN6 Model in Tab. 1

3.2 Evaluation of Different CNN Configurations

The performance of the CNN1 model was evaluated on the test dataset using the MSE and RMSE. RMSE and MSE of the model were evaluated as 0.004353 °C and $1.9648 \cdot 10^{-5}$ $°C^2$, respectively. It should be noted that none of the microstructural configurations and TQC values used during training were also used in the test set. In addition, the model showed an extremely fast evaluation of the temperature increase, making 348 predictions in about 0.4 seconds.

Next, the temperature increase during the last twelve calculation steps, which represent the plastic part of the deformation process, is examined. The comparison of the temperature

increase curves predicted by the ML model with the curves calculated by the FEA for four different hard phase fractions and TQCs is shown in Fig. 10. It is evident that the ML model accurately reproduces the temperature increase obtained by the FEA. Note that the RVE is almost entirely in the elastic state during the first eight calculation steps and the heat generated by plastic dissipation is practically nonexistent. It can also be seen that the temperature increase does not change significantly at different TQCs. Another factor that affects the amount of heat generated is the arrangement of the soft and hard phases within the microstructure, which is studied in Sec 3.3. In the end, it can be noticed that the amount of heat generated depends most strongly on the fraction of the hard phase.

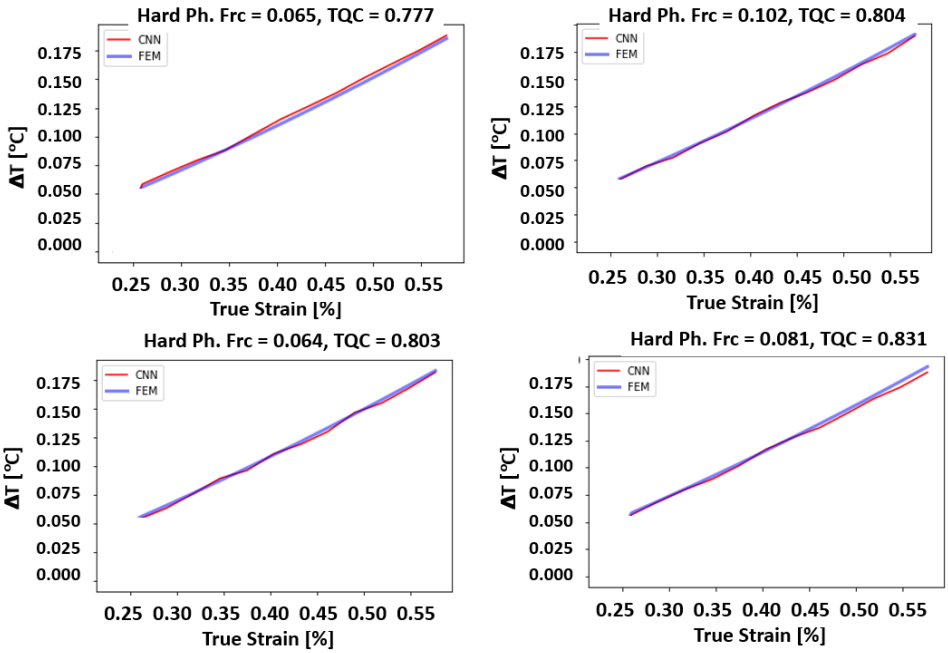

Fig. 10 Comparison of the temperature increase curves as predicted by the ML model with those calculated by the FEM (TQC at the end of deformation is shown)

The performance of the model was also checked for all strain points and compared with the results of the FEA. The comparison of ML and FEA is shown in Fig. 9 (a). Throughout the complete range of temperature increase, the prediction accuracy is very good. It can also be seen that the prediction error for any point in the dataset is less than 6.5%. The errors between the predicted and calculated temperature increase for all strain points were also checked. The maximum error is 0.012 °C, while the average error is $5.87 \cdot 10^{-5}$ °C, while the temperature increase ranges from 0.04 °C to 0.2 °C.

The prediction error of the ML model is also analyzed for different TQC values for all strain points, Figs. 11, 12. Again, it was found that the temperature increase is only weakly related to the TQC value. The prediction error is also not dependent on the TQC value. The error plots for all strain points are shown in Figs. 13, 14 for the different hard phase fractions. The general trend of a temperature increase with increasing hard phase fraction

could be observed. This means that the stiffer RVE configuration increases the amount of plastic strain and the temperature increase. The effect of stiffness is more influential for the temperature increase than TQC value. Nevertheless, a stochastic distribution can be recognized, which is related to the random distribution of the soft and hard phases within the RVE. Consequently, the highest temperature increase is obtained at the fractions of 0.085, 0.118 and 0.135, as this random configuration generates the largest plastic strain. Thus, as already mentioned, the 3D CNN directly captures the phase fractions, the arrangement of the phases and the TQC value. Again, the prediction error does not appear to depend on the hard phase fraction, while the hard phase fraction tends to be the most influential factor for the temperature increase.

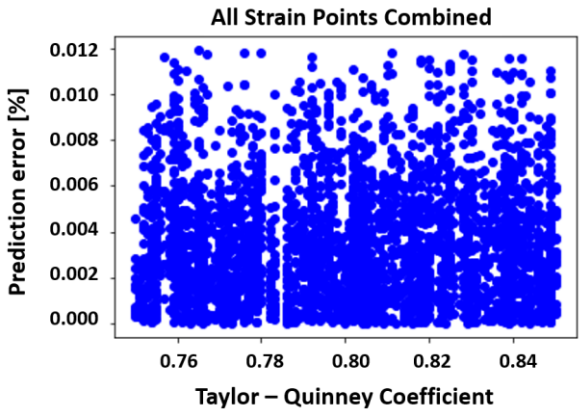

Fig. 11 Prediction error for the given TQC interval of all 12 points of the temperature increase curve – actual TQC (β) for the last strain point is shown

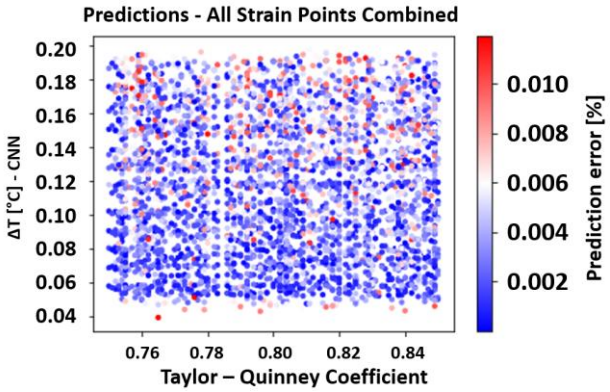

Fig. 12 Predicted and calculated temperature increase for different temperature values of TQC of all 12 points of the temperature increase curve – actual TQC (β) for the last strain point is shown

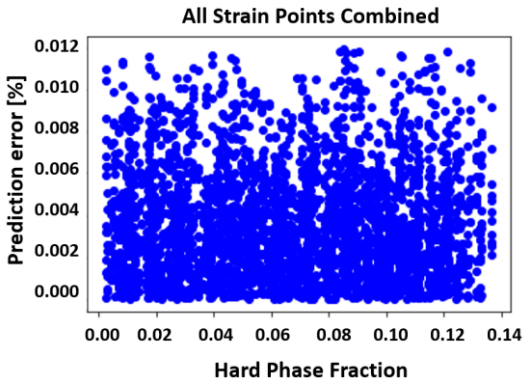

Fig. 13 Prediction error for the hard phase fraction interval of all 12 points of the temperature increase curve

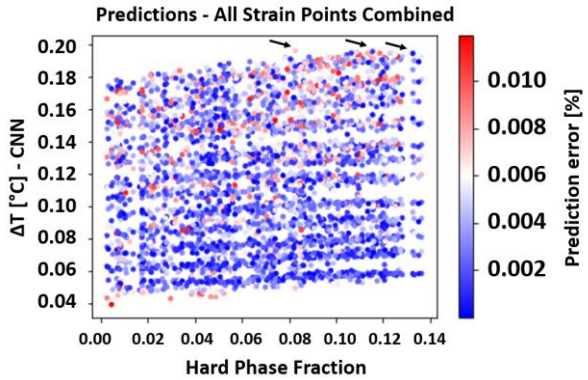

Fig. 14 Predicted and calculated temperature increase for different temperature values of the hard phase fraction of all 12 points of the temperature increase curve - arrows show the highest temperature increase

Finally, this approach is significantly more complex compared to classical approaches that do not use machine learning, e.g. [4]. In the latter, the generated heat is estimated by relying on constant TQC, which is of course only an approximation. The former approach, advocated in the present research is certainly more involved and computationally intensive, at least at start. However, once the neural network is trained, it is fast and can account for differences in microstructure and provide more accurate TQC estimates.

3.3 Influence of the Regularity of Phases Distribution

As shown in Fig. 15, the influence of the variations in distances between the centers of gravity (CoGs) within the different RVEs was investigated. The positions of the CoGs can serve as an indicator for the regularity of the hard phase distribution. A global CoG was calculated for the hard and soft phases of each RVE, with the resulting distances between the CoGs of the soft and hard phases. It should be noted that each strain point is shown in

the graph. Generally, a higher hard phase fractions are associated with a smaller CoG distance. It was also observed that some of the RVEs have significantly higher CoG distances than most of the others. This observation can be traced back to scenarios with very low fractions of hard phase, where a limited number of small particles are dispersed away from the central region of the RVE. This displacement causes the global CoG of the hard phase to deviate significantly from the central point of the RVE, which in turn leads to a larger CoG distance. Recall that the CoG of both phases is generally centered, since phase generation in DREAM.3D is defined via the normal distribution. However, due to the much higher fraction of soft phase in the RVE, the CoG of the soft phase is generally centered and close to the CoG of the RVE. Conversely, an increase in the hard phase fraction generally leads to a reduction in the CoG distance.

When Fig. 15 (b) is evaluated again, a previously drawn conclusion can be confirmed once more: a larger temperature increase is associated with a larger fraction of the hard phase. Furthermore, the prediction error also shows no discernible relationship to the CoG distance. Again, it can be seen that the main factor for the temperature increase is the hard phase fraction.

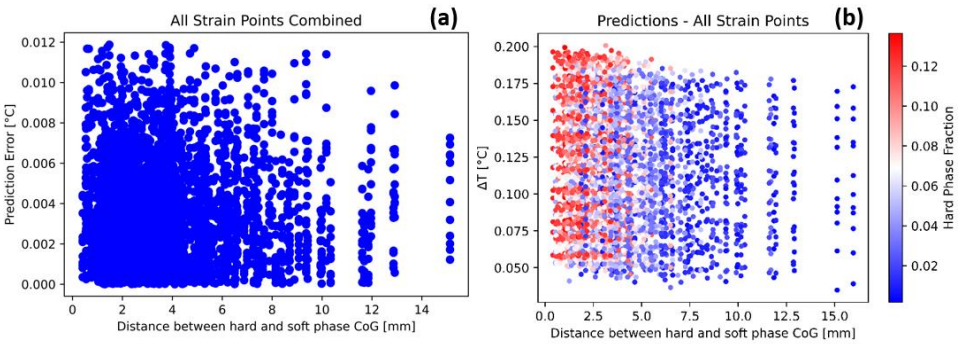

Fig. 15 Distance between center of gravity of hard and soft phase of all 12 points of the temperature increase curve: (a) with prediction error and (b) with temperature increase

Furthermore, a study was carried out to examine the impact of the standard deviation of the distances between all components in the soft and hard phase of the RVEs from the center, see Figs. 16, 17. The main objective is to demonstrate the variability in the dispersion of the data has an influence on the results. From Figs. 16, 17 it is clear that there is no significant correlation between the error and the fraction of the hard phase as well as the soft phase, regardless of whether the standard deviation values are higher or lower.

It can be observed that most of the values of the standard deviation of the soft phase are mostly concentrated in a narrow range around the value of 3.62. The extreme values of the standard deviations of CoG are 3.71 and 3.45, respectively. For the hard phase, it could be observed in Fig. 17 that the values of the standard deviation are concentrated in a narrow range around the value of 3.5, while the minimum and maximum are 0.25 and 5.80, respectively. This is in line with the previous conclusion that the CoG of the hard phase can vary significantly due to the smaller number of particles and the involved stochastic distribution.

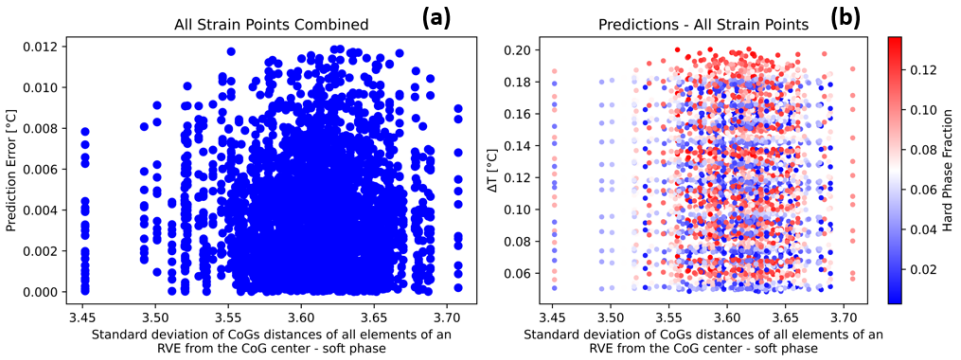

Fig. 16 Standard deviation of center of gravity distances of all elements of an RVE from the center of gravity center for all RVEs of all 12 points of the temperature increase curve – soft phase: (a) with prediction error and (b) with temperature increase

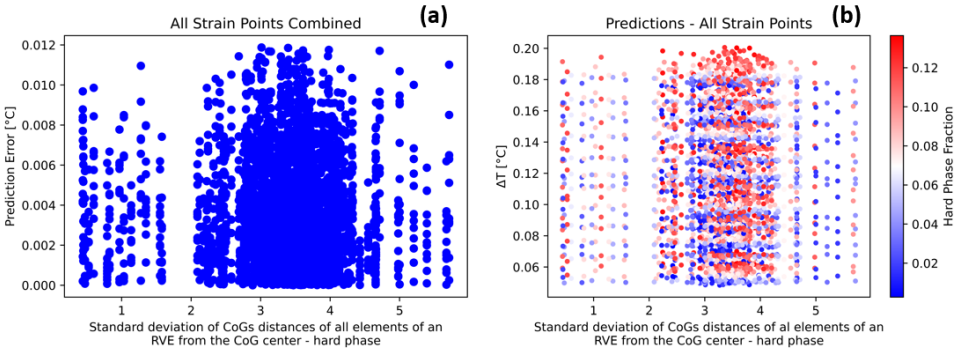

Fig. 17 Standard deviation of center of gravity distances of all elements of an RVE from the center of gravity center for all RVEs of all 12 points of the temperature increase curve– hard phase: (a) with prediction error and (b) with temperature increase

Fig. 18 illustrates the correlation between the variables through a scatterplot representation that includes aspects such as the CoG distances of the hard phase from the center, the standard deviation, the mean distance of the hard and soft phase from the center, the hard phase fractions, and the prediction error. It should be noted that for the calculation of the mean distance, the CoGs of the individual elements within the RVE are calculated for the individual phase and the mean distance is calculated. Upon evaluation of the data, a discernible pattern emerges: The standard deviation of the CoG distances and the mean distance of the hard phase are predominantly concentrated in a narrow range around a certain value. Only a fraction of the RVEs deviate from this specific zone, especially those corresponding to a very small amount of hard phase fractions. This pattern confirms the previous assertion of a normal distribution of phases within the RVE. The CoG distances show a distribution around smaller values, approximately 3 mm, while some RVEs show higher values up to 15 mm. As already explained, these correlate predominantly with smaller fractions of the hard phase.

It should also be again noted that higher values of the temperature increase are primarily associated with higher fractions of the hard phase. Therefore, even though the arrangement of the phases slightly affects the temperature increase and is captured by the ML model, the dominant variable influencing the temperature increase is the hard phase fraction. The error plots are also not correlated with any of the above variables.

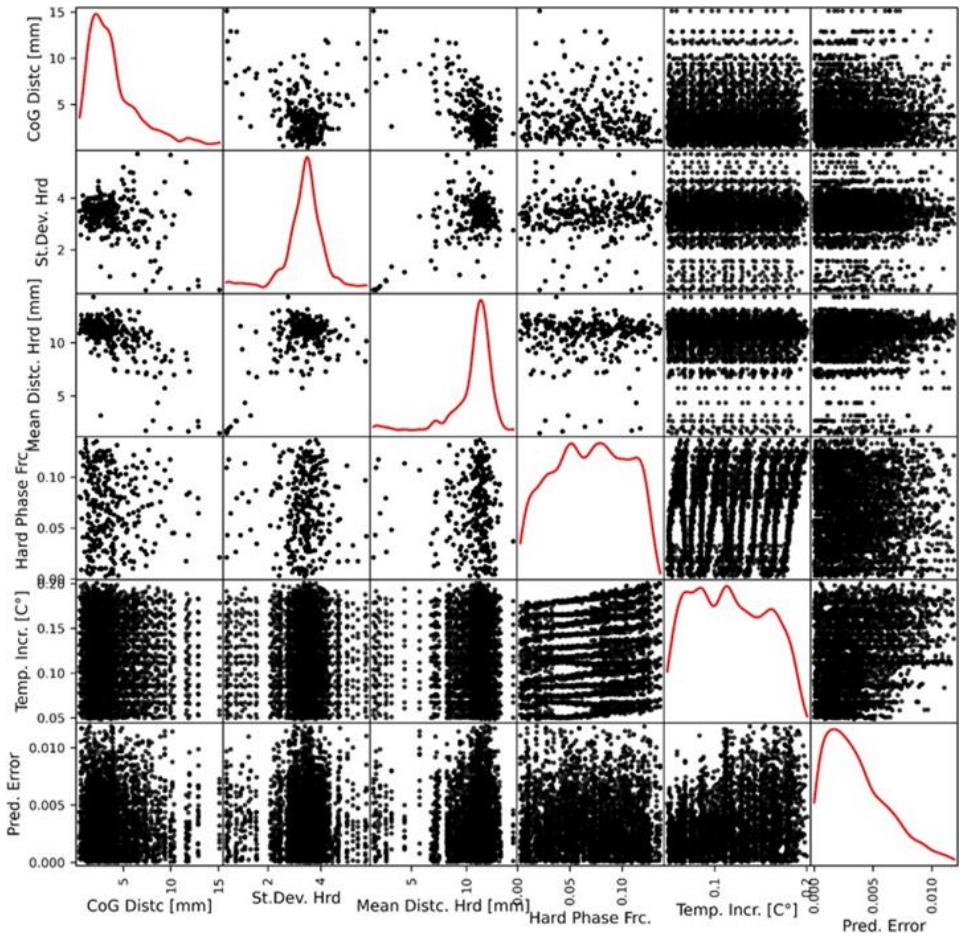

Fig. 18 Scatter matrix of all 12 points of the temperature increase curve of all RVEs – CoG distances of the hard phase from the center, standard deviation, mean distance of hard and soft phase from the center, hard phase fractions, and the prediction error

4. CONCLUSION

The current study has shown the potential for the use of ML methods in predicting temperature changes resulting from plastic deformation of materials using SPR. In particular, more classical approaches based on FEA are more efficient when a smaller

number of microstructure configurations are to be analyzed. These usually rely on Eq. (1). However, if one wants to gain a deeper insight into the problem, the number of configurations increases, which makes the investigation more difficult. An ML model needs to be trained, but once it is trained, it is very efficient. This in turn makes it possible to analyze a larger number of simulations, as the data can be easily processed and the evaluation time is short. This can be useful, for example, when optimizing the microstructure to obtain a specific property since the procedure requires a large number of evaluations.

The study examined the impact of various factors, including the fraction of hard phase, the TQC value and the regularity of phase distributions, on the extent of heat generation. The results revealed that within the TQC range investigated, the TQC value had little influence on the outcomes. In contrast, the fraction of the hard phase emerged as the most influential factor in determining the extent of the temperature increase.

The effect of the regularity of the phase distribution was also investigated. Specifically, the CoG distances between the hard and soft phases and the standard deviation of the CoGs for the soft and hard phases were examined. However, it was found that none of these values correlated with the prediction error or the magnitude of the temperature increase. On the other hand, the investigation reconfirmed that the fraction of the hard phase and the microstructure configuration remains the most influential factor contributing to temperature increase.

Moreover, the CNN model developed in the current study exhibits a high computational efficiency, enabling it to generate nearly 350 predictions within 0.4 s. The model performed very well, achieving an RMSE of 0.004353 and an MSE of $1.9648 \cdot 10^{-5}$. The maximum prediction error of the model is 0.012 °C, while the average error is $5.87 \cdot 10^{-5}$ °C within the temperature range of 0.04 °C to 0.2 °C.

The scope of this study is limited to examining the boundary conditions that simulate a macroscopic uniaxial tensile test, specifically focusing on cases with relatively small plastic deformations. However, it would be worthwhile to explore this phenomenon further by investigating alternative boundary conditions, such as shear loading, and considering larger plastic deformations, provided that sufficient computational resources are accessible.

Acknowledgement: *This work was supported in part by the University of Rijeka under project number uniri-iskusni-tehnic-23-37 and in part by the Croatian Science Foundation under project IP-2019-04-4703. All the support is gratefully acknowledged.*

REFERENCES

1. Čanađija, M., 2023, *Thermomechanics of Solids and Structures: Physical Mechanisms, Continuum Mechanics, and Applications*, Elsevier.
2. Behrens, B.A., Chugreev, A., Bohne, F., Lorenz, R., 2019, *Approach for modelling the Taylor-Quinney coefficient of high strength steels*, Procedia Manufacturing, 29, pp. 464-471.
3. Čanađija, M., Mosler, J., 2016, *A variational formulation for thermomechanically coupled low cycle fatigue at finite strains*, International Journal of Solids and Structures, 100, pp. 388-398.
4. Čanađija, M., Brnić, J., 2004, *Associative coupled thermoplasticity at finite strain with temperature-dependent material parameters*, International Journal of Plasticity, 20(10), pp. 1851-1874.
5. Rusinek, A., Klepaczko, J., 2009, *Experiments on heat generated during plastic deformation and stored energy for TRIP steels*, Materials & Design, 30(1), pp. 35-48.

6. Rittel, D., 1999, *On the conversion of plastic work to heat during high strain rate deformation of glassy polymers*, Mechanics of Materials, 31(2), pp. 131–139.
7. Kositski, R., Mordehai, D., 2021, *Employing molecular dynamics to shed light on the microstructural origins of the Taylor-Quinney coefficient*, Acta Materialia, 205, 116511.
8. Zaera, R., Rodríguez-Martínez, J.A., Rittel, D., 2013, *On the Taylor–Quinney coefficient in dynamically phase transforming materials. application to 304 stainless steel*, International Journal of Plasticity, 40, pp. 185–201.
9. Čanađija, M., Munjas, N., Brnić, J., 2019, *Thermodynamically consistent homogenization in finite strain thermoplasticity*, International Journal for Multiscale Computational Engineering, 17(2), pp. 99-120.
10. DeCost, B.L., Holm, E.A., 2015, *A computer vision approach for automated analysis and classification of microstructural image data*, Computational Materials Science, 110, pp. 126–133.
11. Azimi, S.M., Britz, D., Engstler, M., Fritz, M., Mücklich, F., 2018, *Advanced steel microstructural classification by deep learning methods*, Scientific Reports, 8(1), 2128.
12. DeCost, B.L., Lei, B., Francis, T., Holm, E.A., 2019, *High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel*, Microscopy and Microanalysis, 25(1), 2129.
13. Ferguson, M.K., Ronay, A., Lee, Y.T.T., Law, K. H., 2018, *Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning*, Smart and Sustainable Manufacturing Systems, 2(1), pp. 137 - 164.
14. Nikolić, F., Štajduhar, I., Čanađija, M., 2021, *Casting microstructure inspection using computer vision: Dendrite spacing in aluminum alloys*, Metals, 11(5), 756.
15. Nikolić, F., Štajduhar, I., Čanađija, M., 2022, *Casting defects detection in aluminum alloys using deep learning: A classification approach*, International Journal of Metalcasting, 17, pp. 386–398.
16. Yucel, B., Yucel, S., Ray, A., Duprez, L., Kalidindi, S.R., 2020, *Mining the correlations between optical micrographs and mechanical properties of cold-rolled HSLA steels using machine learning approaches*, Integrating Materials and Manufacturing Innovation, 9(3), pp. 240–256.
17. Tagimalek, H., Maraki, M.R., Mahmoodi, M., Azargoman, M., 2022, *A hybrid SVM-RVM algorithm to mechanical properties in the friction stir welding process*, Journal of Applied and Computational Mechanics, 8(1), pp. 36–47.
18. Milićević, I., Popović, M., Dučić, N., Vujičić, V., Stepanić, P., Marinković, D., Ćojbašić, Ž., 2024, *Improving the mechanical characteristics of the 3D printing objects using hybrid machine learning approach*, Facta Universitatis Series Mechanical Engineering, doi: 10.22190/FUME220429036M.
19. Zarezadeh, A., Shishesaz, M.R., Ravanavard, M., Ghobadi, M., Zareipour, F., Mahdavian, M., 2023, *Electrochemical and mechanical properties of Ni/g-C3N4 nanocomposite coatings with enhanced corrosion protective properties: A case study for modelling the corrosion resistance by ANN and ANFIS models*, Journal of Applied and Computational Mechanics, 9(3), pp. 590-606.
20. Li, X., Zhang, Y., Zhao, H., Burkhart, C., Brinson, L.C., Chen, W., 2018, *A transfer learning approach for microstructure reconstruction and structure-property predictions*, Scientific Reports, 8, 13461.
21. Jung, J., Yoon, J.I., Park, H.K., Kim, J.Y., Kim, H.S., 2019, *An efficient machine learning approach to establish structure-property linkages*, Computational Materials Science, 156, pp. 17–25.
22. Yang, Z., Yabansu, Y.C., Al-Bahrani, R., Liao, W.K., Choudhary, A.N., Kalidindi, S.R., Agrawal, A., 2018, *Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets*, Computational Materials Science, 151, pp. 278–287.
23. Cecen, A., Dai, H., Yabansu, Y.C., Kalidindi, S.R., Song, L., 2018, *Material structure-property linkages using three-dimensional convolutional neural networks*, Acta Materialia, 146, pp. 76–84.
24. Latypov, M.I., Kalidindi, S.R., 2017, *Data-driven reduced order models for effective yield strength and partitioning of strain in multiphase materials*, Journal of Computational Physics, 346, pp. 242–261.
25. King, E., Li, Y., Hu, S., Machorro, E., 2023, *Physics-informed machine-learning model of temperature evolution under solid phase processes*. Computational Mechanics, 72(1), pp. 125-136.
26. Pantalé, O., Mha, P.T., Tongne, A., 2022, *Efficient implementation of non-linear flow law using neural network into the Abaqus Explicit FEM code*. Finite Elements in Analysis and Design, 198, 103647.
27. Zlatić, M., Čanađija, M. 2023, *Incompressible rubber thermoelasticity: a neural network approach*. Computational mechanics, 71(5), pp. 895-916.
28. Groeber, M.A., Jackson, M.A., 2014, *DREAM.3D: a digital representation environment for the analysis of microstructure in 3D*, Integrating Materials and Manufacturing Innovation, 3(1), pp. 56–72.
29. Nikolić, F., Čanađija, M., 2023, *Deep learning of temperature – dependent stress – strain hardening curves*, Comptes Rendus. Mécanique, 351, pp. 151–170.
30. Li, X., Liu, Z., Cui, S., Luo, C., Li, C., Zhuang, Z., 2019, *Predicting the effective mechanical property of heterogeneous materials by image based modelling and deep learning*, Computer Methods in Applied Mechanics and Engineering, 347, pp. 735–753.

31. Wang, Y., Zhang, M., Lin, A., Iyer, A., Prasad, A. S., Li, X., Zhang, Y., Schadler, L., Chen, W., Brinson, L., 2020, *Mining structure–property relationships in polymer nanocomposites using data driven finite element analysis and multi-task convolutional neural networks*, Molecular Systems Design & Engineering, 5, pp. 962–975.
32. Liu, Z., Wu, C., Koishi, M., 2019, *Transfer learning of deep material network for seamless structure–property predictions*, Computational Mechanics, 64(2), pp. 451-465.
33. Balasivanandha Prabu, S., Karunamoorthy, L., 2008, *Microstructure based finite element analysis of failure prediction in particle-reinforced metal-matrix composite*, Journal of Materials Processing Technology, 207(1), pp. 53–62.
34. Phillion, A.B., Cockcroft, S.L., Lee, P.D., 2009, *Predicting the constitutive behavior of semi-solids via a direct finite element simulation: application to AA5182*, Modelling and Simulation in Materials Science and Engineering, 17(5), 055011.
35. Kim, K., Forest, B., Geringer, J., 2011, *Two-dimensional finite element simulation of fracture and fatigue behaviours of alumina microstructures for hip prosthesis*, Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 225(12), pp. 1158–1168.
36. Patel, S., Vaish, R., Sinha, N., Bowen, C., 2014, *Finite element analysis of the microstructure of AlN-TiN composites*, Strain, 50 (3), pp. 250–261.
37. Wei, J., Chu, X., Sun, X.Y., Xu, K., Deng, H. X., Chen, J., Wei, Z., Lei, M., 2019, *Machine learning in materials science*, InfoMat, 1(3), pp. 338–358.
38. Herriott, C., Spear, A.D., 2020, *Predicting microstructure-dependent mechanical properties in additively manufactured metals with machine- and deep-learning methods*, Computational Materials Science, 175, 109599.
39. Ragone, M., Yurkiv, V., Song, B., Ramsubramanian, A., Shahbazian-Yassar, R., Mashayek, F., 2020, *Atomic column heights detection in metallic nanoparticles using deep convolutional learning*, Computational Materials Science, 180, 109722.
40. Barakbayeva, T., Demirci, F.M., 2023, *Fully automatic CNN design with inception and RESNET blocks*, Neural Computing and Applications, 35, pp. 1569–1580.
41. Bandyopadhyay, M., 2021, *Multi-stack hybrid CNN with non-monotonic activation functions for hyperspectral satellite image classification*, Neural Computing and Applications, 33(21), pp. 14809–14822.
42. Nour, M., Öztürk, Ş., Polat, K., 2021, *A novel classification framework using multiple bandwidth method with optimized CNN for brain - computer interfaces with EEG-FNIRS signals*, Neural Computing and Applications, 33, pp. 15815–15829.
43. Leon-Medina, J.X., Anaya, M., Tibaduiza, D.A., Pozo, F., 2021, *Manifold learning algorithms applied to structural damage classification*, Journal of Applied and Computational Mechanics, 7(Special Issue), pp. 1158-1166.