Material Forming - ESAFORM 2024 Materials Research Forum LLC
Materials Research Proceedings 41 (2024) 2299-2308 https://doi.org/10.21741/9781644903131-253

Fast prediction of the material displacement in open die forging using neural networks

JAGTAP Nikhil Vijay[1,a*], REINISCH Niklas[1,b] and BAILLY David[1,c]

[1]Institute of Metal Forming, RWTH Aachen University, Intzestraße 10, 52072 Aachen, Germany

[a]nikhil.jagtap@ibf.rwth-aachen.de, [b]niklas.reinisch@ibf.rwth-aachen.de, [c]david.bailly@ibf.rwth-aachen.de

Keywords: Open Die Forging, Neural Networks, Data-Driven Modeling, FEA

Abstract. This paper presents a data-driven approach to predict the material displacement in open die forging using neural networks. Training data for different process parameters and workpiece geometries is generated using finite element simulations. A neural network architecture is designed that takes the process parameters and the coordinates of a point in the geometry as inputs and outputs the displacement of that point after the deformation. This is systematically implemented for open die forging, using relevant process information. The neural network model is trained and tested on various FEA-simulations for different process parameters and shows good accuracy and generalization. The model is also able to simulate multiple strokes of a single pass in a fast and efficient way. It is demonstrated how the neural network model can enable building a digital material shadow of open die forging processes. The advantages and limitations of the approach are then further discussed.

Introduction

A complete open die forging process is comprised of incrementally forming the workpiece in between saddles using multiple strokes. This can not only produce desired geometries but also improve material properties. The cast ingots that are the input to the forging often suffer from inherent undesirable microstructure including inhomogeneous, often large, grain size and orientation. Furthermore, segregation as well as cavities and pores could be the result of the casting process leading to insufficient quality of the casted workpiece. It is possible to remove these defects and produce finer microstructure using open die forging.

The flexibility offered by open die forging is exploited to manufacture a variety of semi-finished products like disks, rings, rods, shafts, hollow shafts, flanged shafts, crankshafts, and housings. However, these applications involve multiple manufacturing steps after forging to produce the final product. The workpiece geometry and the properties are transformed in each of this step. Integrating the relevant parameters like material properties between all these production steps is crucial for enabling knowledge-based decision-making in between the steps, quality assurance and long-term predicting e.g. the life of the final product. The large volume of data generated during each production step can be systematically shared and transformed throughout the process chain in a Digital Material Shadow (DMS) [1]. Even though, a lot of the material defects are located in the core region of a casted ingot, functional areas in most of the mentioned products are located mainly on the surface. Hence, the equivalent strain and geometry evolution outside the core is also essential for creating meaningful DMS. Additionally, the ability to simulate in short time can enable to trace every workpiece, thus capturing the deviations and inaccuracies in the individual process. This is especially important for highly loaded products like drive shafts.

Various methods have been used to calculate the strain and temperature evolution along with other parameters by simulating a forging process. Finite element analysis (FEA) is used for most numerical simulations e.g. for the macroscale analysis of the forging process in 2D and 3D [2]. FEA-simulations can be accurate but take larger amount of time and computation power to be able to do so. One current alternative is the use of state-of-the-art fast models that calculate important workpiece parameters quickly. Beyond others, Recker et al. [3] presented a combined approach that can calculate equivalent strain as well as temperature and austenite grain size in the core of the open die forging workpiece. For equivalent strain, the model use Siemers approximation [4] whereas for grain size, the Johnson-Mehl-Avrami-Kolmogorov (JMAK) model is used [5]. The change in the global geometry of the workpiece can be calculated using e.g. the spread equations given by Tomlinson and Stringer [6] to superimpose these results. These models provide fast results contrary to the FEA-simulations and thus find applications in optimization and online monitoring [7]. Here, the expected process results can be visualized and enhanced by online process adaption. The results are only limited to the core fiber (Fig. 1), which is an imaginary longitudinal line passing through the center of the workpiece.

Besides full black box approaches, advanced data driven methods also offer ways to integrate the knowledge of the known physics with systematically collected data. By doing this the data efficiency of the models can be increased, which is especially useful for use cases with limited amount of data like open die forging. In this study, neural networks are used, which can be understood as highly nonlinear function approximator inspired by biological neurons [8]. According to the universal approximation theorem, neural networks can be capable of approximating any function [9]. Knapp [10] used neural networks to predict the spread in open die forging which is essential for optimal process control. Experimental data was used to train and test this model. Lee [11] used convolutional neural networks to design the preform geometry of axisymmetric closed-die forgings to improve strain distribution. Chan [12] used FEA to train a neural network for predicting parameters like forming load and effective stress for critical design parameters to evaluate the quality of the punch die design.

In this paper, the application of neural networks is studied for simulating the local displacement in 2D cross-section of the forging workpiece. Since various process results like equivalent strain can be inferred from the local displacement, this should enable the creation of a DMS in the future. Various process parameters like bite ratio (bite length b/ height h) and height reduction are used to define a process in open die forging and they directly influence the local deformation. These process parameters thus form an input to the neural networks suiting for a variety of pass schedules. Finally, the accuracy of the neural network model is compared with the FEA results.

Data Collection and Feature Study

In a systematic implementation of machine learning using neural networks, the solution is highly dependent on the model architecture as well as the data the model is trained with. Both of these components can be modeled to convey the physical aspects of the manufacturing process [13]. The insights from the process can be embedded in the data through the data collection methodology and feature study. The data should represent the complete domain without any bias. The domain can be defined by identifying the key variables in open die forging that affect the displacement. Hence, a short discussion of the main influences on the displacement follows. The open die forging process consists of multiple forging passes, where the workpiece is typically rotated around the longitudinal axis mostly based on the target geometry. During each pass, the workpiece is incrementally forged usually from one end of the workpiece to the other through multiple strokes. Each stroke deforms the workpiece locally in between the saddles. All the strokes in a given pass are characterized by process parameters like bite ratio and height reduction. All the passes and their corresponding process parameters are listed in the pass schedule.

According to Knapp [10], the spread in the workpiece is influenced by the ratio of height to width (here referred to as the aspect ratio α of the lateral cross-section),

$$\alpha_0 = \frac{h_0}{w_0} \quad (1)$$

the relative bite length (or bite ratio x_b),

$$x_b = \frac{b}{h_0} \tag{2}$$

relative deformation (or height reduction ε_h),

$$\varepsilon_h = \frac{h_0 - h_1}{h_0} \tag{3}$$

temperature and steel grade. As shown in Fig. 1, a long workpiece with rectangular cross-section with initial width w_0, height h_0 and length l_0 is considered. The longitudinal direction of the workpiece is along z-axis and the press moves along y-axis. The core fibre passes through the core of the workpiece in z-direction. The material spreads freely in the x-direction during each stroke, thus changing the aspect ratio of the rectangular cross-section.

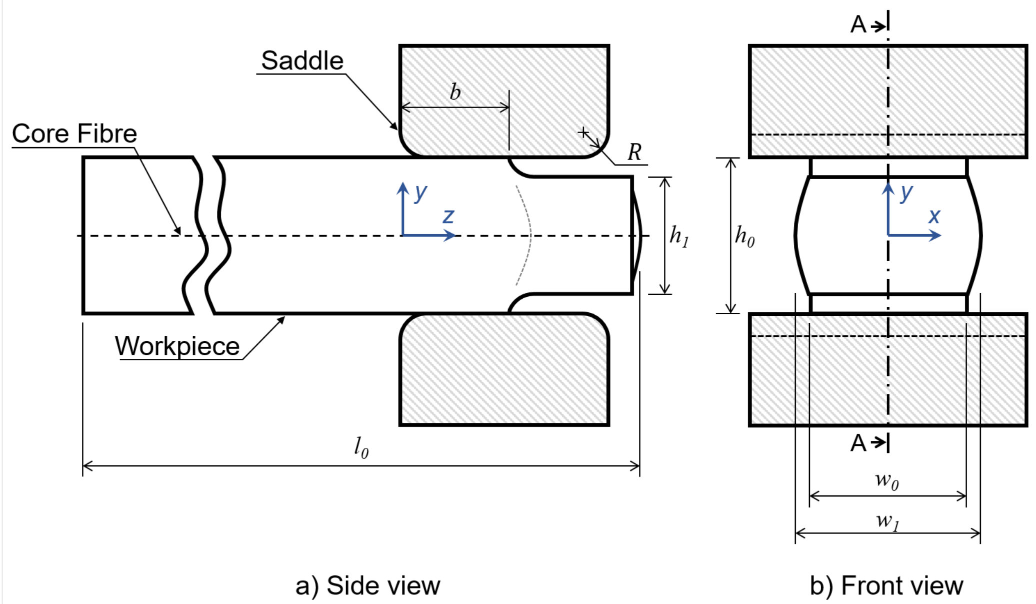

Fig. 1. A typical open die forging workpiece and saddle in side view (a) and front view (b) along with the respective coordinate systems.

These three parameters, aspect ratio, bite ratio and height reduction given in Eq. 1, Eq. 2 and Eq. 3 form the variables of the study due to their influence on the spread as well as process description. The domain of each parameter or the *variable space* is defined in Table 1. These ranges are most commonly observed in practical forging processes. Each point in this variable space refers to a single *process* that is characterized by corresponding parameters.

Table 1. Important process parameters and their corresponding ranges that form the variable space.

Variables	**Range [-]**
$\boldsymbol{\alpha_0}$ - aspect ratio	0.9 to 1.5
$\boldsymbol{x_b}$ - bite ratio	0.3 to 0.9
$\boldsymbol{\varepsilon_h}$ - height reduction	5% to 26%

To begin with, two datasets are gathered using FEA for the complete variable space defined in the Table 1. The first, respectively the training dataset, is used for training and validation of the neural networks. The sampling of the processes from the variable space was kept random yet ensuring sufficient representation of each variable using *latin hypercube sampling* (LHS). Altogether, 200 processes were sampled from the variable space for the training dataset. The second, respectively the test dataset is a collection of 54 processes from the variable space that are not part of the training dataset. The performance of the trained model is determined based on how accurate it predicts on the unseen test data. The accuracy of the model on this unseen dataset indicates the ability to generalize the underlying patterns in the data.

Each process sampled above is simulated for the corresponding values of aspect ratio, bite ratio and height reduction in a FEA software, *Simufact Forging 15.0*. Addressing the requirement of high accuracy data, a template simulation was setup in Simufact. The parameters that affect the quality and the correctness of the results are defined in this template. Mesh convergence was performed to select suitable mesh size optimized to the application. Symmetry planes were not used to allow more data points that can be augmented by mirroring. For simplicity, the material grade and starting temperature are fixed in this study. A commonly used alloy steel 42CrMo4 (1.7225) is chosen in this study. The furnace or starting temperature is set to 1100°C. The height and width of the workpiece depends on the aspect ratio, however for the square aspect ratio ($\alpha_0 = 1$) the height of the workpiece is 100 mm. Each simulation includes three forging strokes to reach a steady deformation state, since the first two strokes exhibit deviating deformation due to their proximity to the ingots end. The simulation time increments are automatically determined by Simufact and the results of those at the end of each sub step, like at the end of the feed, end of the stroke and end of the backstroke, are saved.

A software program is developed in MATLAB, which creates copies of the template and writes the corresponding process parameters from the variable space to a new simulation. Later, it runs the simulation on the designated machine and the results are stored. A data handling program was written to manage this distributed database while extracting, transforming and loading the data on demand so that it can be used by neural networks for respectively training, validation and testing.

Displacement Prediction Model based on Neural Networks

A process inspired architecture of the model, that incorporates features arising from the forging process, can bear meaningful results. Thus, the architecture is systematically developed by first studying the data. This method is highly iterative and the final model is a result of multiple adaptions in the data, model architecture and the hyper parameters in the training process. After each adaption, the network is tested for certain criteria derived from their importance to the forging process like error in the surface nodes. These criteria are mentioned in the next chapter. The following analysis of the features and model architecture refer to the final model.

This study is limited to a 2D plane in the direction of the press and passing through the core fiber (plane A-A in Fig. 1), which in these simulations is the yz-plane. Nevertheless, within the data collection, 3D FEA was used and the 2D cross-sections are extracted from 3D results. Fig. 2 shows the displacement in each y and z-directions for the third stroke in an example simulation (α_0=1.4, x_b=0.75, ε_h=17%). The deformation zone can be distinguished from the rest of the workpiece in this figure. Some nodes in the center of the deformation zone are not displaced in the z-direction, mainly due to friction. However, those outside show almost constant displacement indicating the increase in length in both the positive and negative z-direction. On the contrary, there is no displacement in y-direction in this region outside the deformation zone and a large movement of nodes due to the movement of press in y-direction.

The center of deformation changes based on the stroke number, bite ratio and height reduction. All the deformation zones can be positioned uniformly by moving the origin of the coordinate system to the left edge of the saddle. This way, all the nodes under the press will lie in positive z-direction starting from the origin (Fig. 1 and Fig. 2), enabling the model a reliable guess of the general material flow.

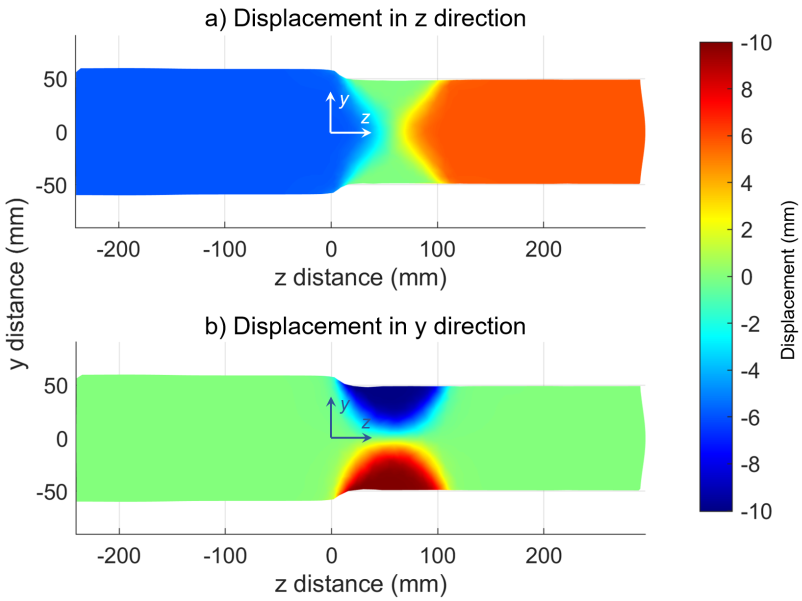

Fig. 2. An example of a forging stroke (α_0=1.4, x_b=0.75, ε_h=17%) and resulting displacements of nodes in y and z-direction respectively.

Several neural network architectures are suitable for this type of problem. However, a simple feed-forward artificial neural network was used for this study. This type of network has a fixed number of inputs and outputs and is well suited to approximate non-linearity in the functions. Other commonly used image-based alternatives limit the inputs to a fixed square grid of pixels. But in the case of this problem, restricting workpieces of different lengths to a fixed image size will make it resolution dependent.

In this study, MATLAB is used to program the neural networks. The input, output and the structure of the neural network is determined by defining reference configuration and deformed configuration at t=t_0 and t respectively. This is similar to the formulation of non-linear problems in continuum mechanics. In this case, deformed configuration is the workpiece after the force is applied and the material is deformed for the given stroke, whereas the reference configuration is the workpiece before the application of the force. Let P be a position vector to an arbitrary point in the workpiece such that in point P(x_0, y_0) becomes P_1(x_1, y_1) in the deformed configuration. The displacement of any point between the two configurations can be calculated as,

$$\mathrm{u} = P_1 - P_0 \tag{4}$$

Let f be any function that maps any point in the deformed configuration to its displacement. Additionally, this function depends on the process parameters that determine the deformation and therefore the displacement of any point in the given configuration. Thus, this gives the position of the point in the deformed configuration:

$$P_1 = P - f(P, \alpha_0, \mathrm{x_b}, \varepsilon_\mathrm{h}) \tag{5}$$

The function f is replaced by a neural network (final adaption consists of 16 layers with a total of 4800 neurons), as shown in Fig. 3. It can predict the coordinates in the deformed configuration from the process parameters and a given coordinate in the reference configuration. The output vector is the two components of the displacement which is then used to calculate the corresponding nodes in the deformed configuration. This network can run multiple times for each node in the single geometry to thus create a result for the complete body, since the input vector is only defined for a single node. Finally, the resulting model can be trained and tested to predict the deformed part geometry after the forging stroke, independent of the mesh size used.

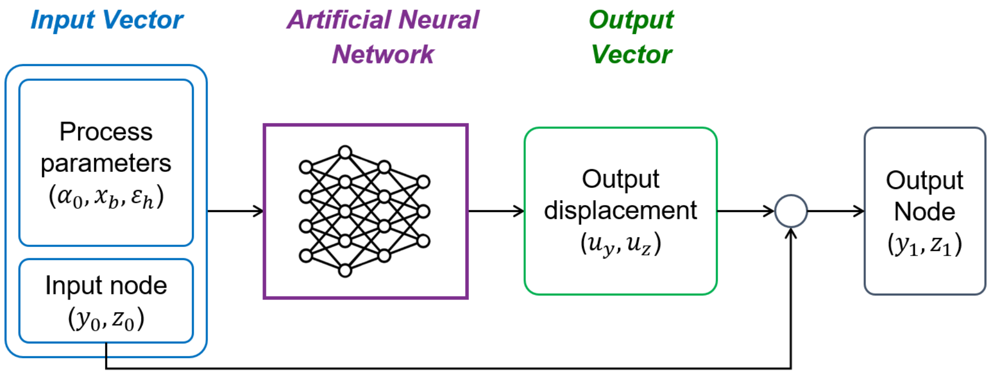

Fig. 3. Flowchart of the model used for predicting displacement using neural networks.

Results and Discussion

The process of training neural networks involves calculating the loss by comparing the model results with the training data, in this case FEA results. The loss is then used to modify the *learnables* or *weights and biases* in the neurons. The training is performed in small batches of data. These batches are shuffled after every epoch, which is a complete iteration through the whole dataset. The training was performed on the 200 simulation results in the training dataset, which consisted of a total of 806,821 nodes or input vectors. It took around 20 epochs or 42,500 iterations on the 16 layer neural network with 4800 weights and biases (approximately 70 min training time).

The test data, which is collected separately, is a full factorial sampling of the variable space resulting in 54 simulations that are not a part of the training data. Testing the neural networks is essential not just to evaluate the network on unseen data but also to see how the model performs. Firstly, the error in the predicted coordinates and the actual coordinates from the FEA results is analyzed in the deformation zone. Secondly, this error is analyzed for the surface nodes separately since they form the boundary of the geometry, and hence define the workpiece volume that needs to obey the volume constancy. Finally, the root mean square error (RMSE) is analyzed for the complete geometry and analyzed for each process in the test dataset. Additionally, the time required for the simulation is also analyzed.

Fig. 4 shows the model output for an exemplary process from the test dataset corresponding to a large height reduction and a large bite ratio. The input to the model (Fig. 4 a) is the configuration before the third stroke. Here, the workpiece is aligned to the origin based on the bite length. With around 4000 nodes in the geometry, the model predicts the deformed configuration after the third stroke. The output (Fig. 4 b) of the model is the position of each node as predicted by the model, then overlaid with the error in position as compared with the FEA results. The predictions in the regions outside the deformation zone are predicted with a higher accuracy. A systematic error is visible in the deformation zone in the region corresponding to a large deformation. However, the overall RMSE for this geometry is 0.16 mm with all the nodes under 2 mm error.

Additionally, the errors in the longitudinal surface are plotted in Fig. 4 c). In the deformation zone, the error is relatively high with peaks in certain nodes. These nodes are consistently present in the region with large deformation, especially in the region under the radius of the

Materials Research Proceedings 41 (2024) 2299-2308 https://doi.org/10.21741/9781644903131-253

saddle.

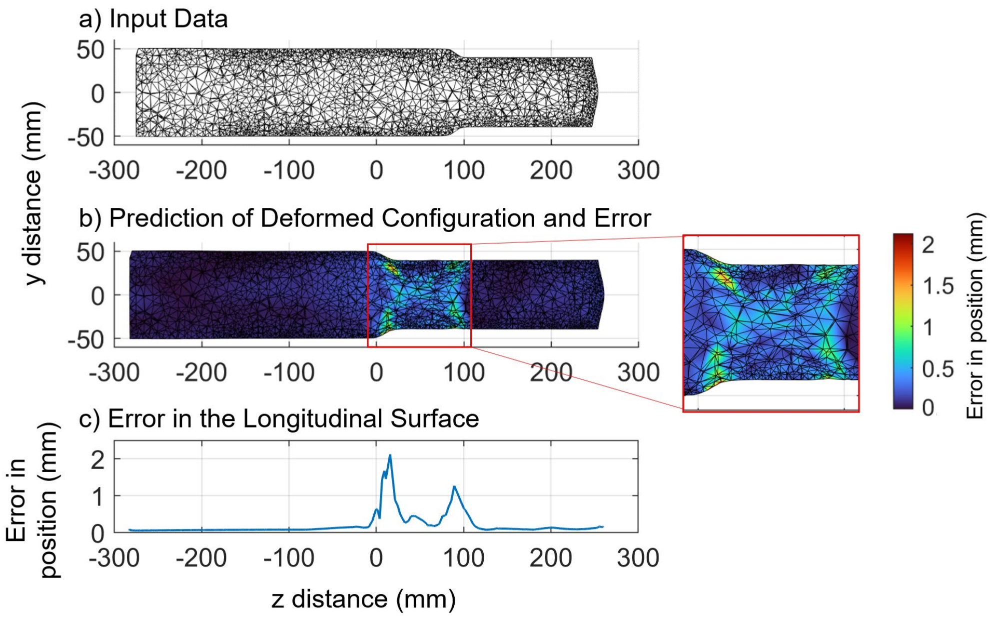

Fig. 4. Prediction of third stroke of forging (b) and its comparison with FEA result for an example process in the test dataset (α_0=1, x_b=0.75, ε_h=21%).

Fig. 5 shows the results for the different predictions in the test dataset corresponding to aspect ratio 1. Fig. 5 a) is a heat map of RMSE between the predicted and actual position for all the nodes of geometry in the given simulation. The maximum error in each of these simulations is shown in Fig. 5 b). The results presented here correspond to the third stroke, similar to the previous results. The maximum errors can be weakly correlated to the RMSE. Overall, RMSE for the complete test dataset is 0.486 mm. Predicting a single stroke for around 4000 nodes takes approximately 0.06 seconds (Intel Xeon E-2286G, NVIDIA Quadro P400).

It can be seen in Fig. 5 that the simulations with a smaller height reduction show consistently higher errors. This was also seen for other aspect ratios and it is mainly due to the smaller overall displacement values as compared to the size of the workpiece in these combinations. In addition, there are certain combinations with comparatively higher errors, which was also seen in the previous adaptions. For some of these outliers, no systematic explanation from the forging process can be found, except for the reasons explained earlier. It is challenging to control certain behaviors of the model due to a variety of factors that affect the prediction, such as the distribution of the data, randomness during the training process, hyperparameter values like learning rate and number of training epochs. The large number of neurons that form the neural network and their corresponding weights and biases make them less explainable (so-called "black box") and require thorough testing. Furthermore, due to the multidimensional input and output relationships, it is not possible to test the model for all possible combinations of variables. These challenges apply to such solutions.

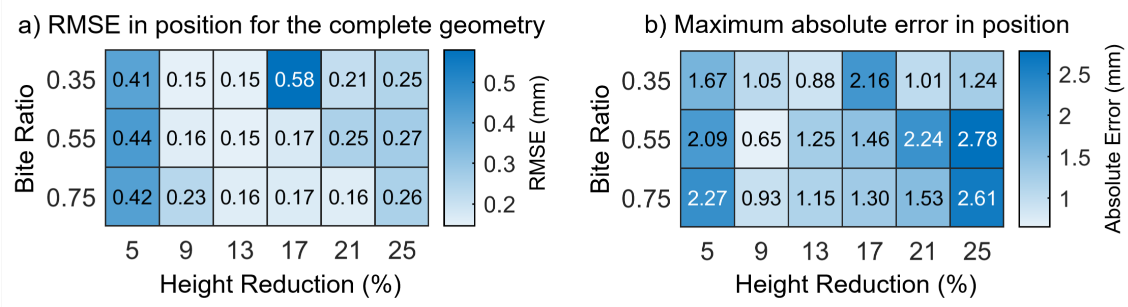

Fig. 5. Evaluation of the model over various processes corresponding to aspect ratio 1.

Overall, the model successfully predicts the displacement and thus the deformed configuration of the forging stroke. The average error is well under 1% relative to the height of the workpiece. The majority of the data (around 70%) in any given simulation is outside the deformation zone and does not show any complex deformation patterns. In spite of this bias in the data, the model is able to predict deformation zone with a fair accuracy. The z-coordinates of the nodes that are undergoing deformation always lie in the same region, the positive z-direction and closer to the origin, making the corresponding features distinguishable from the remaining data.

The model has shown an excellent generalization in the variable space and the 2D space. It predicts good accuracy on the test data which belongs to the processes in variable space that are not a part of the training data. Additionally, the nodes that are tested also do not completely coincide with the training data. Any point in this 2D space can be predicted successfully using this model thus making it spatially independent. The spatial independence can be further demonstrated by using the same model, trained on a coarse mesh, to predict a geometry with finer mesh. A randomly selected process (α_0=1, x_b=0.66, ε_h=23%) is simulated for a coarse and a finer mesh. A corresponding coarse mesh results in 3851 data points as compared to 10834 data points in the finer mesh. The results indicate a similar accuracy with the maximum error increasing from 2.28 mm to 2.50 mm. The overall RMSE for coarse mesh is 0.173 mm whereas for the fine mesh is 0.163 mm. This indicates spatial generalization as the nodes that are not part of the training data (which had coarse mesh) are still predicted with a comparable accuracy by the neural networks. The node-by-node design of the neural networks make it independent of the number of nodes in input model and the network learns the underlying patterns in the 2D space.

Finally, even though the model is trained on the third forging stroke, the repositioning of the training data makes the model virtually independent of the stroke number. That is, the model can predict any stroke, as long as the reference configuration of the geometry is correctly aligned to the press (similar to Fig. 1) in regards to the bite length of the next stroke. This can be demonstrated by taking an arbitrary set of points that form a rectangular geometry before the first stroke of forging. The model predicts the deformed configuration corresponding to the first stroke. This geometry can be translated in the z-direction to align to the press as mentioned before. The resulting geometry is the reference configuration of the second stroke. In this way, a geometry can be incrementally deformed using the same model for predicting each stroke.

Fig. 6 shows the incremental simulation of forging strokes using the developed 2D displacement model as described above for a randomly selected process from the test data (α_0=1, x_b=0.55, ε_h=17%). Each predicted stroke is compared with the FEA result in order to evaluate the accuracy.

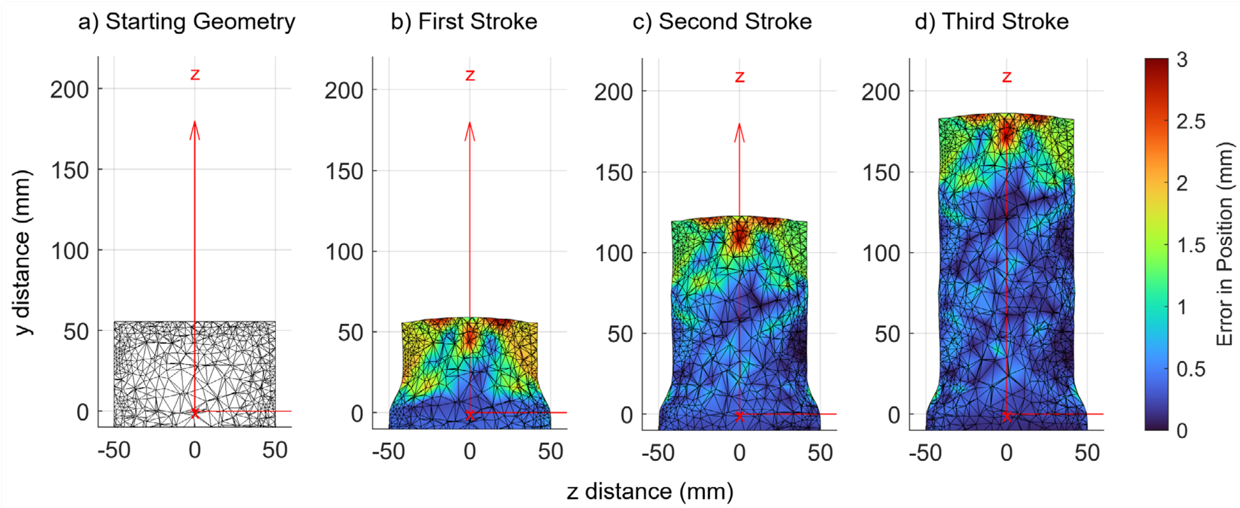

Fig. 5. An example process that is predicted incrementally to simulate a pass. The results are compared to a corresponding FEA result and the error in the position is overlaid on the prediction.

After three strokes, the RMSE of the entire prediction is 0.61 mm thus showing a good generalization of the strokes and ability to simulate passes without a finite element model. The maximum error in position is however 3 mm. The first stroke has relatively poor accuracy since it deforms non-homogeneously and is not a part of the training data. However, the remaining strokes exhibit a closer prediction as the 90% of all the nodes have error in position lower than 1.16 mm. This ability of the model to generalize the different strokes does not directly arise from the neural network architecture but from the feature study and the preprocessing of the data.

Conclusion and Outlook

The neural network model presented in this paper can predict the 2D local displacement in the geometry for any open die forging process in the applied variable space with an overall RMSE of 0.486 mm for a workpiece with around 100 mm in height. This approach predicts a single node at a time and shows a good spatial generalization thus making it resolution independent. Along with its capability to generalize the patterns in the variable space, the model is also able to predict the 2D material flow of any stroke for the given pass. Although the model is data driven, the application of known physics in collecting the data, studying the features, building the neural network architecture and evaluating the model, has been beneficial in achieving the found generalization. Finally, the ability of the model to predict the 2D displacement accurately in milli seconds makes it suitable for the use in a digital material shadow and related applications.

In the future it is planned to improve the existing 2D model e.g. by incorporating the effects of other variables like friction and temperature. Furthermore, an extension of the 2D model to predict 3D material flow seems possible. This extension enables tracing the workpiece geometry throughout complete forging processes including e.g. intermediate workpiece rotations. Finally, the model can be made more explainable and robust by embedding boundary conditions and the underlying physical conditions into the neural networks using custom loss functions and concepts like physics informed neural networks.

Acknowledgement

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612

References

[1] F. Becker, P. Bibow, M. Dalibor, A. Gannouni, V. Hahn, C. Hopmann, M. Jarke, I. Koren, M. Kröger, J. Lipp, J. Maibaum, J. Michael, B. Rumpe, P. Sapel, N. Schäfer, G. J. Schmitz, G. Schuh, and A. Wortmann, "A Conceptual Model for Digital Shadows in Industry and Its Application," in Conceptual Modeling: 40th International Conference, ER 2021, Virtual Event, October 18–21, 2021, Proceedings, 2021, pp. 271–281.

[2] J. Yanagimoto, D. Banabic, M. Banu, and L. Madej, "Simulation of metal forming – Visualization of invisible phenomena in the digital era," CIRP Annals, vol. 71, no. 2, pp. 599–622, 2022, https://doi.org/10.1016/j.cirp.2022.05.007

[3] D. Recker, M. Franzke, and G. Hirt, "Fast models for online-optimization during open die forging," CIRP Annals, vol. 60, no. 1, pp. 295–298, 2011, https://doi.org/10.1016/j.cirp.2011.03.142.

[4] E. Siemer, "Qualitätsoptimierende Prozeßsteuerung des Reckschmiedens," 1987.

[5] C. M. Sellars, "Modelling microstructural development during hot rolling," Materials Science and Technology, vol. 6, no. 11, pp. 1072–1081, 1990, https://doi.org/10.1179/mst.1990.6.11.1072

[6] A. Tomlinson, and J. Stringer, "Spread and elongation in flat tool forging," Journal of the Iron and Steel Institute, 1959.

[7] D. Rosenstock, D. Recker, G. Hirt, K. J. Steingießer, R. Rech, B. Gehrmann, and R. Lamm, "Application of a Fast Calculation Model for the Process Monitoring of Open Die Forging Processes," KEM, 554-557, pp. 248–263, 2013, https://doi.org/10.4028/www.scientific.net/KEM.554-557.248

[8] K. Mehrotra, Elements of artificial neural networks. Cambridge, Mass: MIT Press, 1997.

[9] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," Neural Networks, vol. 2, no. 5, pp. 359–366, 1989, https://doi.org/10.1016/0893-6080(89)90020-8

[10] M. Knap, G. Kugler, H. Palkowski, and R. Turk, "Prediction of Material Spreading in Hot Open-Die Forging," steel research international, vol. 75, no. 6, pp. 405–410, 2004, https://doi.org/10.1002/srin.200405787

[11] S. Lee, K. Kim, and N. Kim, "A Preform Design Approach for Uniform Strain Distribution in Forging Processes Based on Convolutional Neural Network," J. Manuf. Sci. Eng, vol. 144, no. 12, 2022, https://doi.org/10.1115/1.4054904

[12] W. L. Chan, M. W. Fu, and J. Lu, "An integrated FEM and ANN methodology for metal-formed product design," Engineering Applications of Artificial Intelligence, vol. 21, no. 8, pp. 1170–1181, 2008, https://doi.org/10.1016/j.engappai.2008.04.001

[13] M. Mozaffar, S. Liao, X. Xie, S. Saha, C. Park, J. Cao, W. K. Liu, and Z. Gan, "Mechanistic artificial intelligence (mechanistic-AI) for modeling, design, and control of advanced manufacturing processes: Current state and perspectives," Journal of Materials Processing Technology, vol. 302, p. 117485, 2022, https://doi.org/10.1016/j.jmatprotec.2021.117485