



Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Full Length Article

Spatiotemporal prediction of microstructure evolution with predictive recurrent neural network

Amir Abbas Kazemzadeh Farizhandi^a, Mahmood Mamivand^{b,*}^a Computer Science Department, Boise State University, United States^b Department of Mechanical and Biomedical Engineering, Boise State University, United States

ARTICLE INFO

Keywords:

Deep learning
Spatiotemporal prediction
Material microstructure evolution
Predictive Recurrent Neural Network (PredRNN)
Phase field method

ABSTRACT

Prediction of microstructure evolution during material processing is essential to control the material properties. Simulation tools for microstructure evolution prediction based on physical concepts are computationally expensive and time-consuming. Therefore, they are not practical when either there is an urgent need for microstructure morphology during the process or there is a need to generate big microstructure datasets. Essentially, microstructure evolution prediction is a spatiotemporal sequence prediction problem, where the prediction of material microstructure is difficult due to different process histories and chemistry. We propose a Predictive Recurrent Neural Network (PredRNN) model for the microstructure prediction, which extends the inner-layer transition function of memory states in LSTMs to spatiotemporal memory flow. As a case study, we used a dataset from spinodal decomposition simulation of FeCrCo alloy created by the phase-field method for training and predicting future microstructures by previous observations. The results show that the trained network predicts quantitatively accurate microstructure morphologies while it is several orders of magnitude faster than the phase field method.

1. Introduction

The process-structure-property relationships of engineered materials are directly impacted by material microstructures, which are mesoscale structural elements that operate as an essential link between atomistic building components and macroscopic qualities. One of the pillars of contemporary materials research is the ability to manage the evolution of the materials' microstructure while it is being processed or used, including common phenomena like solidification, solid-state phase transitions, and grain growth. Therefore, a key objective of computational materials design has been comprehending and forecasting microstructure evolution. On the other hand, the integration of big data and ML in materials science has greatly increased our understanding of materials and has opened up new avenues for research and innovation. Available and open-source database is a big challenge in using artificial intelligence in material science, particularly for microstructure processing. Creating big data in a wide range of processing conditions is a gap in material design by ML that should be addressed.

Simulations of microstructure evolution frequently rely on coarse-grained models, such as partial differential equations (PDEs) which

are used in the phase-field techniques [1], because they can represent time and length scales that are far larger than those that can be captured by atomistic models. A wide range of significant evolutionary mesoscale processes, including grain development and coarsening, solidification, thin-film deposition, dislocation dynamics, vesicle formation in biological membranes, and crack propagation, have all been fully described using the phase-field method [2–4]. However, there are some significant problems with this strategy as well. First off, PDE-based microstructure simulations are still relatively expensive. The stability of numerical techniques that use explicit time integration for nonlinear PDEs sets stringent upper bounds on the smallest time-step size in the temporal dimension. Similarly, implicit time-integration techniques manage longer time steps by adding more inner iteration loops at each step. Furthermore, despite the fact that in theory controlling PDEs can be inferred from the underlying thermodynamic and kinetic considerations, actual PDE identification, parametrization, and validation take a significant amount of work. The evolution principles may not be fully understood or be too complex to be characterized by tractable PDEs for difficult or less well-studied materials.

Currently, the efforts to reduce computational costs have mostly

* Corresponding author.

E-mail address: mahmoodmamivand@boisestate.edu (M. Mamivand).<https://doi.org/10.1016/j.commatsci.2023.112110>

Received 8 November 2022; Received in revised form 9 February 2023; Accepted 1 March 2023

Available online 16 March 2023

0927-0256/© 2023 Elsevier B.V. All rights reserved.

concentrated on utilizing high-performance computer architectures [5] and sophisticated numerical techniques [6], or on merging machine learning algorithms with simulations based on microstructures [7–9]. Leading studies, for instance, have developed surrogate models using a variety of techniques, such as Green's function solution [10], Bayesian optimization [11], a combination of dimensionality reduction and autoregressive Gaussian processes [12], convolutional autoencoder and decoder [13], or integrating a history-dependent machine-learning method with a statistically representative, low dimensional description of the microstructure evolution generated directly from phase-field simulations that can quickly predict the evolution of the microstructure from phase-field simulations [14]. The main problem, however, has been to strike a balance between accuracy and computing efficiency, even for these successful systems. For complex, multi-variable phase-field models, for example, precise answers cannot be guaranteed by the computationally effective Green's function solution. In contrast, complex, coupled phase-field equations can be solved using Bayesian optimization techniques, however, at a higher computational cost (although the number of simulations required is kept to a minimum because the Bayesian optimization protocol determines the parameter settings for each subsequent simulation). The capacity of this class of models to predict future values outside of the training set is constrained by the fact that autoregressive models can only forecast microstructural evolution for the values for which they were trained. In other models based on dimensionality reduction methods like principle component analysis (PCA), a large amount of information is ignored, which will sacrifice accuracy.

To address the above knowledge gaps, in this study we develop a network based on Predictive Recurrent Neural Network-based (PredRNN) [15] to predict the 2D microstructure evolution over time. In recent years, the usage of ML algorithms in materials research has increased significantly [16,17]. They have been used in a variety of ways, including the discovery of new materials [18–21], the prediction of materials' properties [22–24], the creation of precise and effective potentials for atomistic simulations [25], the analysis and processing of microscopic and spectroscopic data [26–31], the successful inference of a material's properties from a small body of experimental data [32,33], and materials chemistry and processing history prediction from sole micrograph [34,35]. These works include microstructure classification and quantification [26,36], image segmentation [27,28], predictions of microstructure-property relations [37,38], mapping processing-microstructure relations [34,35,39–41], microstructure optimization [42,43], and equilibrium configuration prediction [44] are just a few of the many works that focus on material microstructure and have encouraging results. The datasets used in these investigations are primarily static images of microstructure. This research aims to highlight the significant temporal link between microstructure morphologies along their evolutionary history.

Recurrent Neural Networks (RNNs) [45] are neural networks with hidden memory units that are intended to predict temporal data sequences. RNNs have achieved extensive success in natural language processing [46], speech recognition [47], and computer vision [48] thanks to the development of efficient variations such as the Gated Recurrent Unit (GRU) [49], and the long short-term memory (LSTM) [50], which addresses the vanishing gradient problem during back-propagation. Convolutional neural networks (CNNs) and LSTM have recently been proposed for the predictive learning of spatiotemporal sequences [51].

Yang et al. [52] used the Eidetic 3D LSTM (E3D-LSTM) [53] model to predict evolution phenomena in different processes, such as plane-wave propagation, grain growth, spinodal decomposition, and dendritic crystal growth. Although the results show that the model can predict the evolutionary phenomena precisely, two important issues are not addressed in this study. First, all the used morphologies are constructed artificially for both training and validation. On the other hand, for example, spinodal decomposition occurs in two separate phases: a quick

composition modulation growth phase, followed by a slower coarsening phase, during which the Gibbs-Thomson effect causes a progressive rise in the length scale of the phase-separation pattern. While microstructure morphologies change drastically in these two phases, they trained the RNN to detect system evolution in the second coarsening stage. Therefore, developing a model that can predict morphology evolution in both phases is desired. In this study, we demonstrate that PredRNN can precisely capture all the required features from earlier microstructures to predict long-term microstructures.

2. Methods

2.1. Phase-field modeling

Significant improvements in computational power and advances in numerical methods have made the PF approach a powerful tool for quantitative modeling of the temporal and spatial evolution of material microstructures. Some applications of this method in material processing include martensitic transformation [54], crack propagation [55], grain growth [56], and materials microstructure prediction for properties optimization [57].

The PF method eliminates the need for the system to track each moving boundary by having a finite-width interface that gradually transitions from one composition or phase to another [58]. This essentially leads to the system being modeled as a diffusivity problem. This can be solved by using the nonlinear PDEs of the continuum. Two major PF PDEs show the evolution of various PF variables. One is the Allen-Cahn equation [59] for solving unconserved order parameters (phase domain, particles, etc.), and the other is the Cahn-Hilliard equation [60] for solving conserved order parameters (concentration, etc.).

The constituent elements control the process of phase separation. Hence, the microstructure evolution can be found by tracking the conserved variables, i.e., Fe, Cr, and Co concentration in the isothermal spinodal phase decomposition process. Therefore, our model is governed by the Cahn-Hilliard equation. Equations (1) and (2) are the Cahn-Hilliard equations for the spinodal decomposition of the Fe-Cr-Co ternary system. The PF model for data generation is mainly adopted from Koyama and Onodera's study [61].

$$\frac{\partial c_{Cr}}{\partial t} = \nabla \cdot M_{Cr,Cr} \nabla \frac{\delta F_{tot}}{\delta c_{Cr}} + \nabla \cdot M_{Cr,Co} \nabla \frac{\delta F_{tot}}{\delta c_{Co}} \quad (1)$$

$$\frac{\partial c_{Co}}{\partial t} = \nabla \cdot M_{Co,Cr} \nabla \frac{\delta F_{tot}}{\delta c_{Cr}} + \nabla \cdot M_{Co,Co} \nabla \frac{\delta F_{tot}}{\delta c_{Co}} \quad (2)$$

where c_{Cr} and c_{Co} are the concentrations of Chromium and Cobalt, t is time, F_{tot} is total free energy, and M is the mobility function. The evolution of microstructures is primarily driven by the minimization of the total free energy, F_{tot} of the system. The model is parametrized using a calculation of phase diagram (CALPHAD) data, for details please refer to Ref [34,35]. The Multiphysics Object-Oriented Simulation Environment (MOOSE) is used to solve the nonlinear PDEs. MOOSE is an open-source finite element package developed by the Idaho National Laboratory and is efficient for parallel computing on supercomputers [62]. The combined Cahn-Hilliard equations were solved using the weak form of the residual series of MOOSE's pre-designed Cahn-Hilliard PDEs.

2.2. Training and testing dataset

This study uses the morphology of microstructures of the Fe-Cr-Co ternary alloy in different temperatures and compositions as training and testing datasets. To cover all the ranges of parameters, as mentioned in Table 1, the simulation conditions are designed by the design of experiments. Since the parameters contain chemical compositions and are subject to the constraint that they must sum to one, the mixture design as a design of the experiment method is adopted [63]. Unlike temperature and chemistry, we did not grid the time domain linearly because the

Table 1
Simulation variables and their range of values for database generation.

Simulation variable	Range of values	Grid	
Time (Sec.)	10–1080000	10–3600	50
		3600–36000	500
		36000–360000	5000
		360000–1080000	100,000
Temperature (K)	850–970	10	
Chromium composition (at. %)	0.05–0.9	0.05	
Cobalt composition (at. %)	0.05–0.9	0.05	

microstructure is susceptible to aging time in the early stages of annealing, but as time passes, this sensitivity drops considerably. We generated the microstructures by solving the Cahn-Hilliard PDEs using the MOOSE framework. The simulations were run on different clusters including Boise State University R2 cluster computers [64], Boise State University BORAH [65], and the Extreme Science and Engineering Discovery Environment (XSEDE) (Jetstream2 cluster), which is supported by National Science Foundation (NSF) [66].

After running the simulations, the microstructures were collected from the results. The Fe-based composition microstructure morphologies sequences are utilized to construct the dataset. The length of each sequence is 20 microstructures; the first 10 microstructures are used to predict the future 10 microstructures.

2.3. Spatiotemporal predictive

Prediction of microstructure evolution is a spatiotemporal problem. Different network architectures, which can generally be grouped into three categories: feed-forward models based on CNNs, recurrent models, or a combination of convolution and recurrent networks, the Transformer-based, and flow-based methods, are used to encode different inductive biases into neural networks for spatiotemporal predictive learning [67]. The inductive bias of group invariance over space has been brought into spatiotemporal predictive learning through the use of convolutional layers. For next-frame prediction in Atari games, Oh et al. [68] defined an action-conditioned autoencoder with convolutions. The Cross Convolutional Network, developed by Xue et al. [69], is a probabilistic model that stores motion data as convolutional kernels and learns to predict a likely set of future frames by understanding their conditional distribution. In order to complete the crowd flow prediction challenge, Zhang et al. [70] suggested using CNNs with residual connections. It specifically takes into account the proximity, duration, trend, and external elements that affect how population flows move. Additionally, the convolutional architectures are employed in tandem with the generative adversarial networks (GANs) [71], which successfully lowered the learning process' uncertainty and enhanced the sharpness of the generated frames. Most feed-forward models demonstrate greater parallel computing efficiency on large-scale GPUs compared to recurrent models [72–74]. However, these models generally fail to represent long-term reliance across distant frames since they learn complex state transition functions as combinations of simpler ones by stacking convolutional layers.

Some helpful insights into how to forecast upcoming visual sequences based on historical observations are provided by recent developments in RNNs. In order to forecast future frames in a discrete space of patch clusters, Ranzato et al. [75] built an RNN architecture that was influenced by language modeling. As a remedy for video prediction, Srivastava et al. [76] used a sequence-to-sequence LSTM model from neural machine translation [77]. Later, other approaches to describe temporal uncertainty or the multimodal distribution of future frames conditioned on historical observations have been presented, integrating variational inference with 2D recurrence [78–80]. By arranging 2D recurrent states in hierarchical designs, certain additional

techniques successfully increased the forecast time horizon. The factorization of video information and motion is another area of research, typically using sequence-level characteristics and temporally updated RNN states [81]. The use of optical flows, new adversarial training schemes, relational reasoning between object-centric content and pose vectors, differentiable clustering techniques, amortized inference enlightened by unsupervised image decomposition, and new types of recurrent units constrained by partial differential equations are typical approaches [82–84]. The aforementioned techniques work well for breaking down dynamic visual scenes or understanding the conditional distribution of upcoming frames. To describe the spatiotemporal dynamics in low-dimensional space, they primarily use 2D recurrent networks, which inadvertently results in the loss of visual information in actual circumstances.

Shi et al. [85] created the Convolutional LSTM (ConvLSTM), which substitutes convolutions for matrix multiplication in the recurrent transitions of the original LSTM to combine the benefits of convolutional and recurrent architectures. A schematic of ConvLSTM with a basic LSTM cell is given in Fig. 1. An action-conditioned ConvLSTM network was created by Finn et al. [86] for visual planning and control. Shi et al. [87] coupled convolutions with GRUs and used non-local neural connections to expand the receptive fields of state-to-state transitions. Wang et al. [88] introduced a higher-order convolutional RNN that uses 3D convolutions and temporal self-attention to describe the dynamics and includes a time dimension in each hidden state. Su et al. [89] increased the low-rank tensor factorization-based higher-order ConvLSTMs' computational effectiveness. Convolutional recurrence provides a platform for further research by simultaneously modeling visual appearances and temporal dynamics [90,91]. The spatiotemporal memory flow, a novel convolutional recurrent unit with a pair of decoupled memory cells, and a new training method for sequence-to-sequence predictive learning are all used to enhance the existing architectures for action-free and action-conditioned video prediction in Predictive Recurrent Neural Network (PredRNN) [15].

A network component known as a memory cell is crucial in helping stacked LSTMs solve the vanishing gradient issue seen by RNNs. It can latch the gradients of hidden states inside each LSTM unit during training, preserving important information about the underlying temporal dynamics, according to strong theoretical and empirical evidence. However, the spatiotemporal predictive learning task necessitates a distinct focus on the learned representations in many areas from other tasks of sequential data; therefore, the state transition pathway of LSTM memory cells may not be optimum. First, rather than capturing spatial deformations of visual appearance, most predictive networks for language or speech modeling concentrate on capturing the long-term, non-Markovian features of sequential data [92]. However, both space-time data structures are essential and must be carefully considered in order to forecast future frames. Second, low-level features are less significant to outputs in other supervised tasks using video data, such as action recognition, where high-level semantical features may be informative enough. The stacked LSTMs don't have to maintain fine-grained representations from the bottom up because there are no complex structures of supervision signals. Although the current inner-layer memory transition-based recurrent architecture can be sufficient to capture temporal variations at each level of the network, it might not be the best option for predictive learning, where low-level specifics and high-level semantics of spatiotemporal data are both significant to generating future frames. Wang et al. [15] proposed a new memory prediction framework called PredRNN, which extends the inner-layer transition function of memory states in LSTMs to spatiotemporal memory flow. This framework aims to jointly model the spatial correlations and temporal dynamics at different levels of RNNs. All PredRNN nodes are traversed by the spatiotemporal memory flow in a zigzag pattern of bi-directional hierarchies: A newly created memory cell is used to deliver low-level information from the input to the output at each timestep, and at the top layer, the spatiotemporal memory flow transports the high-

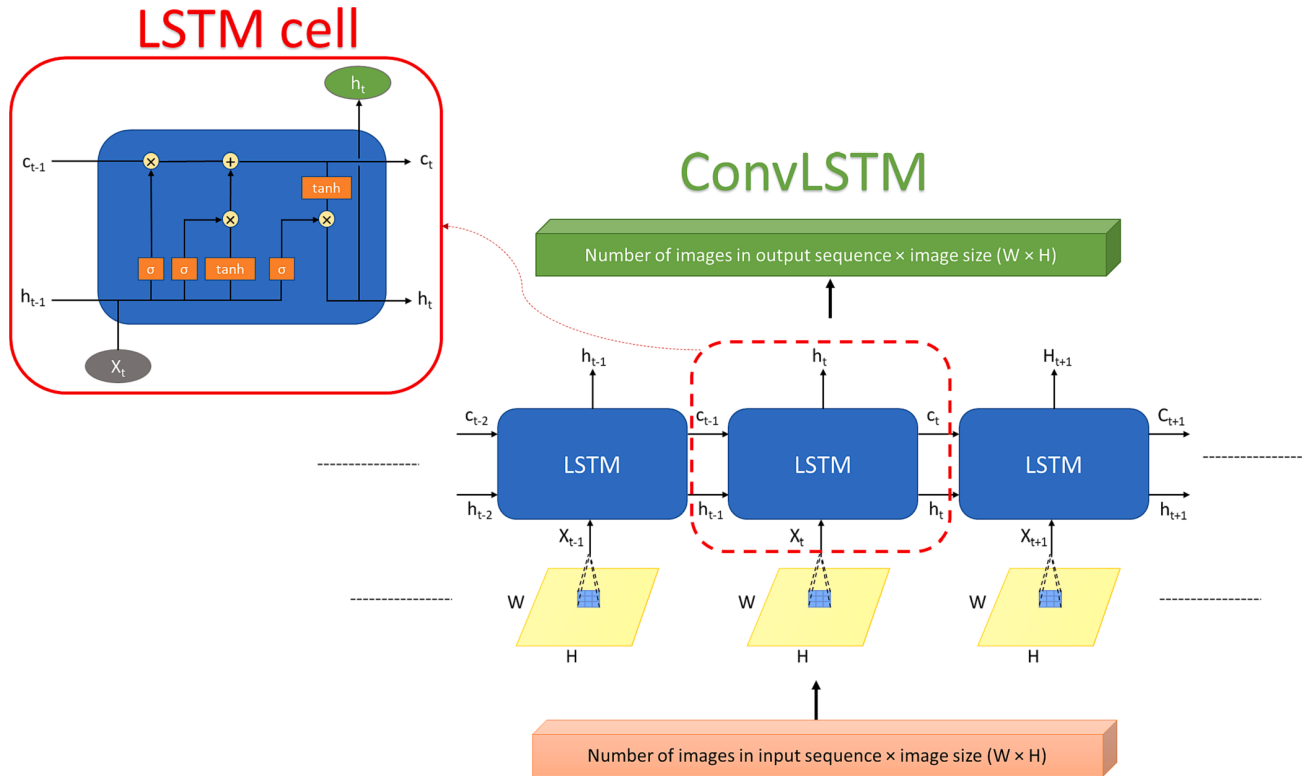


Fig. 1. Structure of ConvLSTM with basic LSTM cell.

level memory state to the bottom layer at the following timestep. The Spatiotemporal LSTM (ST-LSTM), in which the proposed spatiotemporal memory flow interacts with the original, unidirectional memory state of LSTMs, was therefore established as the fundamental building element of PredRNN. It seems that they would require a unified memory mechanism to handle both short-term deformations of spatial details and long-term dynamics if they anticipated a vivid imagination of numerous future images: On the one hand, the network may learn complex transition functions within brief neighborhoods of subsequent frames thanks to the new spatiotemporal memory cell architecture, which also increases the depth of nonlinear neurons across time-adjacent RNN states. Thus, it considerably raises ST-modeling LSTM's capacity for short-term dynamics. To achieve both long-term coherence of concealed states and their fast reaction to short-term dynamics, ST-LSTM, on the other hand, still uses the temporal memory cell of LSTMs and closely combines it with the suggested spatiotemporal memory cell. On five datasets—the Moving MNIST dataset[76], the KTH action dataset[93], a radar echo dataset[94] for precipitation forecasting, the Traffic4Cast dataset[95] of high-resolution traffic flows, and the action-conditioned BAIR dataset [86] with robot-object interactions—the proposed methodology demonstrated state-of-the-art performance. The original paper [15] contains information about the investigation in detail. This study adopts the PredRNN to predict the microstructure evolution quickly and accurately.

3. Results and discussion

3.1. Phase-field modeling for microstructure sequences generation

Following the Simplex-Lattice design, the microstructure sequences are produced by the PF modeling of Fe-Cr-Co spinodal decomposition. For this study, 4,212 phase field simulations (18 different Cr compositions, 18 Co different compositions, and 13 different temperatures) were run based on Table 1. Then, the sequences were constructed based on the

generated time series microstructures in each PF run (125,233 microstructures morphology in total). Some sample microstructure sequences from the PF simulation results are shown in Fig. 2. On a 24 Core CPU, a MOOSE simulation of a 200 nm 2D domains uses about 120 service units (SU) every run. Therefore, it took around 505 k SU to screen the suggested range of temperatures and chemical compositions for microstructure evolution. Fig. 2 shows some samples of microstructure evolution and indicates that the microstructure morphology patterning differs in various chemical compositions and temperatures.

The training dataset can be generated from simulated microstructures. The length of each sequence is 20 microstructures. The first 10 microstructures, which are from the first 30 hr of heat treatment, are used to predict the future 10 microstructures, which have heat treatment time between 50 hr to 300 hr. There are 20,000 sequences for training and 4,000 sequences for testing data. Three different Fe-composition-based microstructure morphology sequences are presented in Fig. 3.

Fig. 3 shows the dataset contains very different evolution sequences in terms of structure. In addition, since the microstructures are selected from both distinct stages of spinodal decomposition, i.e., a fast composition modulation growth stage and a slower coarsening stage, the difference between the input and output sequence is significant, which can be easily recognized in Fig. 3. In this case, the model has a more difficult task in predicting the output sequences.

3.2. Microstructure evolution prediction by PredRNN

20,000 sequences trained the PredRNN to predict the output microstructures. With a mini-batch of 8 sequences, we trained the models using the ADAM optimizer. After 80,000 iterations, the training process is terminated with a learning rate of 10^{-4} . PredRNN typically employs four ST-LSTM layers to balance training effectiveness with prediction quality. We set the size of the convolutional kernels inside the ST-LSTM unit to 5×5 and the number of channels of each hidden state to 128.

As illustrated in Fig. 4, the training loss decreases smoothly with

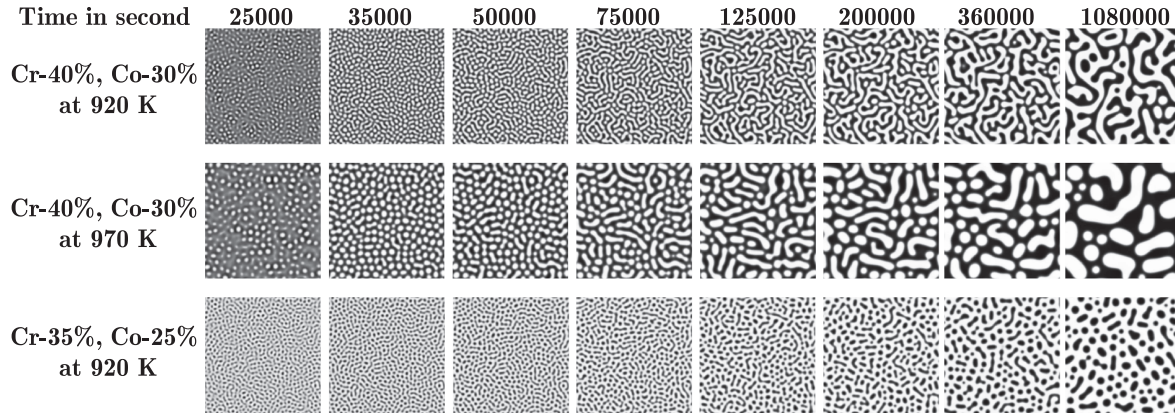


Fig. 2. The Fe-composition base 2D microstructure sequences for different temperatures and chemical compositions produced by the PF method (Compositions are in atomic percent).

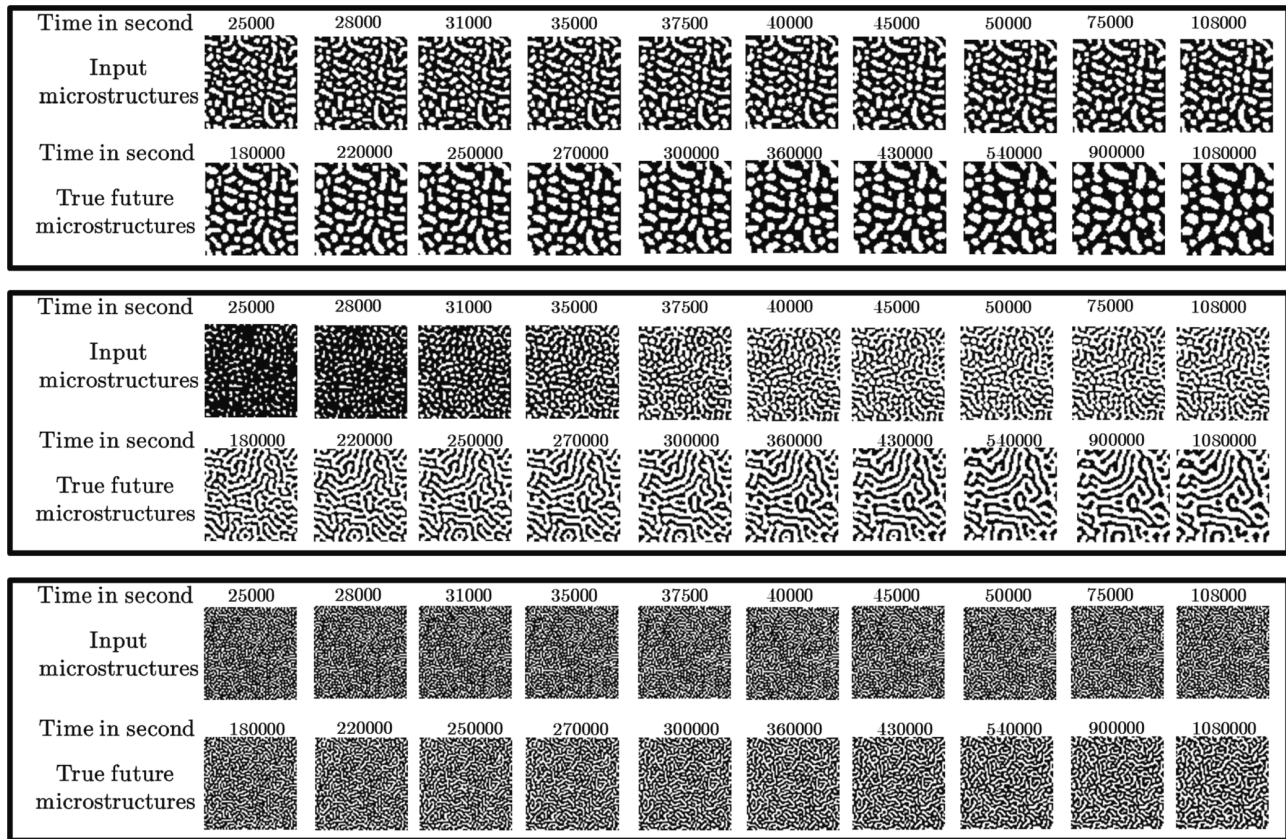


Fig. 3. Three different Fe-composition-based microstructure morphology sequences.

iteration, which indicates that the model's parameters have reached their optimal value globally. In addition, we employ evaluation measures that are frequently used to determine the similarity between two images. The predicted and ground truth microstructures are compared using the Mean Squared Error (MSE), the Peak Signal-to-Noise Ratio (PSNR), the Structural Similarity Index Measure (SSIM), and the Learned Perceptual Image Patch Similarity (LPIPS). In these metrics, MSE estimates the absolute pixel-wise errors, PSNR compares image compression quality, SSIM measures the similarity of structural information within the spatial neighborhoods, and LPIPS is based on deep features and is more in line with human perceptions. Smaller MSE and LPIPS, and higher PSNR and SSIM indicate more similarity between images.

After training, test sequences are used to compute MSE, LPIPS, PSNR, and SSIM; the average values for each iteration are given in Fig. 5. The results demonstrate that all the metrics improve with iteration to reach almost stability. It proves that the model learns from the data and can train the hyperparameters.

Fig. 6 displays three randomly selected samples from the test set for a qualitative comparison. The left microstructures of the dashed line are the input frames, the right ones in the top row are the ground truth of output microstructures, and the bottom row shows the PredRNN prediction. The microstructures produced by PredRNN predict clear images, meaning it can be confident of future variations. In addition, we can see that the predicted sequence is close to the ground truth

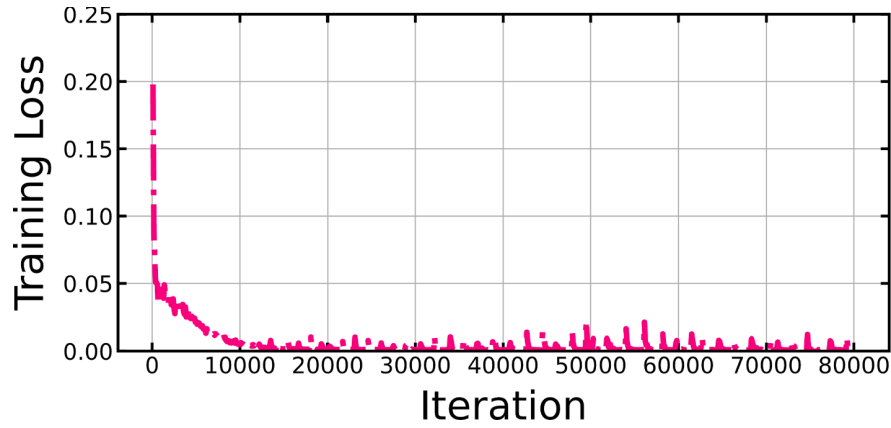


Fig. 4. Training loss per iteration indicates model convergence after 10,000 iterations.

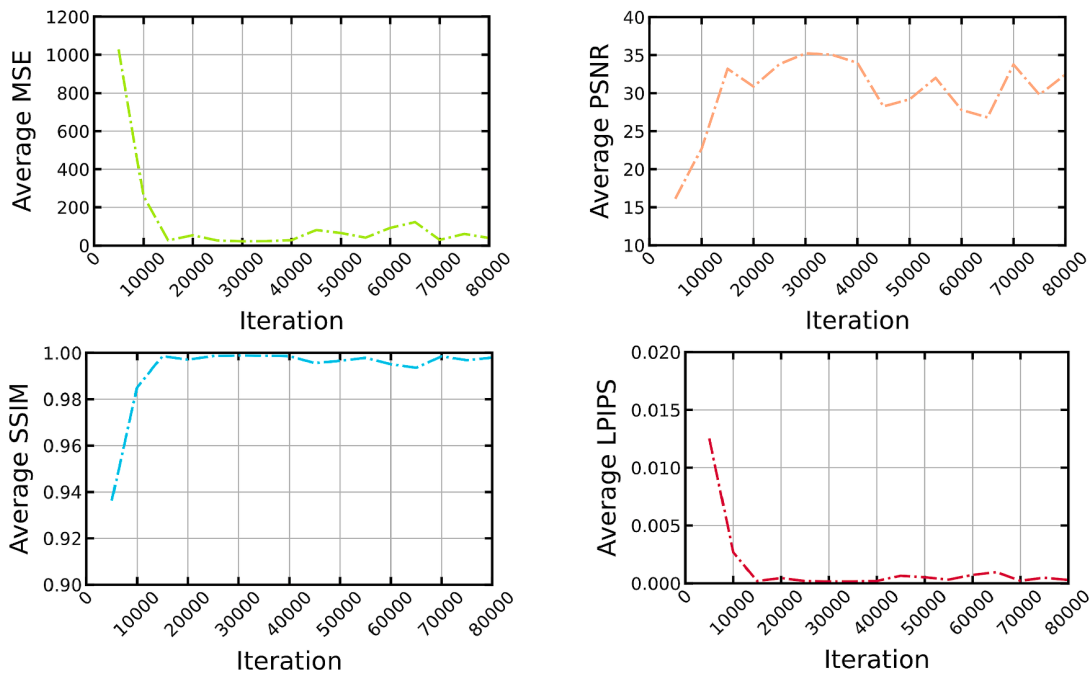


Fig. 5. Average MSE, PSNR, SSIM, and LPIPS for testing sequences during training per each iteration.

sequence.

3.3. Trained model performance on the microstructure evolution prediction during time

Model performance on frames prediction during time is one of the key parameters in spatiotemporal models' evaluation [96,97]. Basically, the prediction of earlier frames is easier than long-term prediction because of the similarity with the input sequence. Fig. 7 provides the corresponding frame-wise comparisons between the predicted by the final PredRNN model and ground truth microstructures for test sequences. The average values of metrics show that the model can predict all the microstructures with reasonable accuracy. On the other hand, the model has higher accuracy in the prediction of the first frames than the last ones, as MSE and LPIPS increase and PSNR and SSIM decrease from frame 1 to 10.

For quantitative comparison of long-term and short-term predictions, similarity/error metrics for the test set produced by the final PredRNN model are given in Fig. 8. The results show that PredRNN

prediction for short-term cases is more accurate than long-term predictions. These results seem reasonable because there is a stronger correlation between the first microstructures from the output sequence and the input sequence. However, in general, the predictions for long-term cases also have good agreement with the ground truth. It proves that the PredRNN can predict the microstructure evolution reasonably well.

3.4. Trained model inference performance in future microstructures prediction

The time it takes to calculate the model's outputs as a function of the inputs is known as the inference speed. The model's response time is crucial in many applications, especially those requiring real-time data [98]. Since this study aims to develop a deep network to predict the microstructure evolution quickly and accurately, the model inference performance is a critical factor. Therefore, the trained model performance is compared with the simulation on a reference computer. Since MOOSE can only run with the CPU, we used the same resource for the

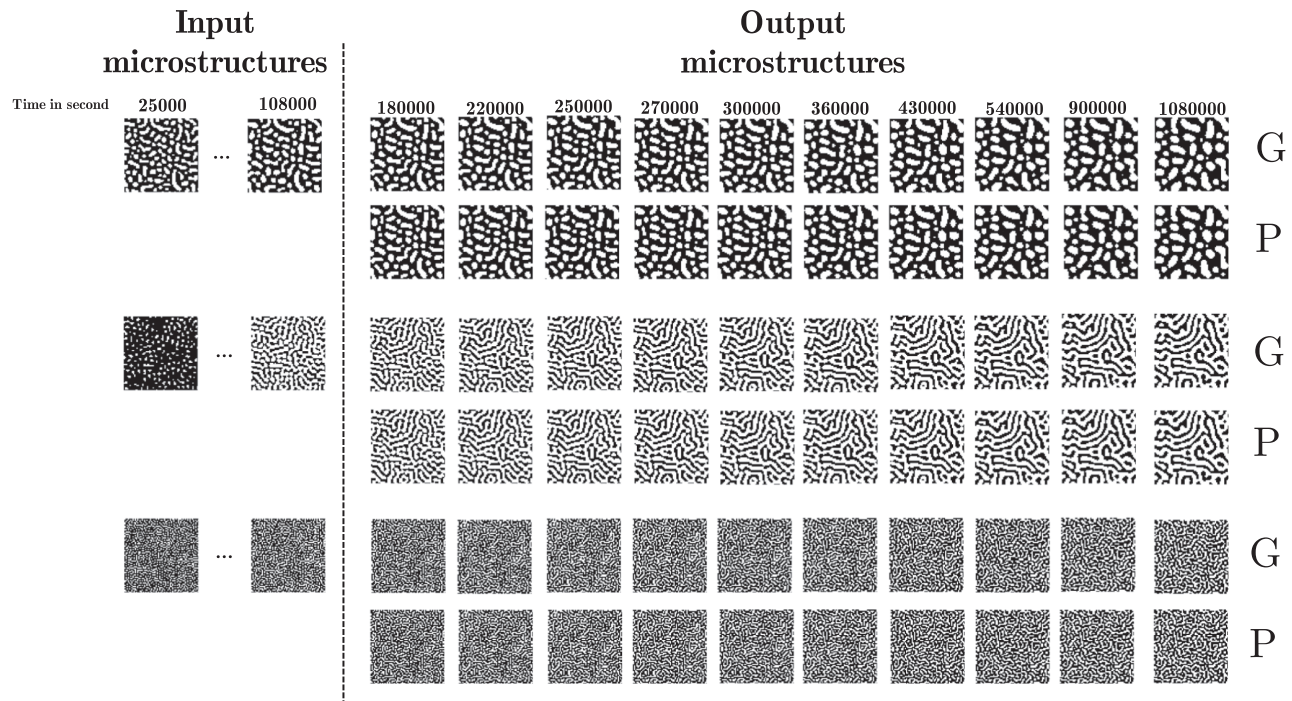


Fig. 6. Frame-wise results on the three randomly selected samples from the test set produced by the final PredRNN model (predictions (P) vs. ground truth (G)).

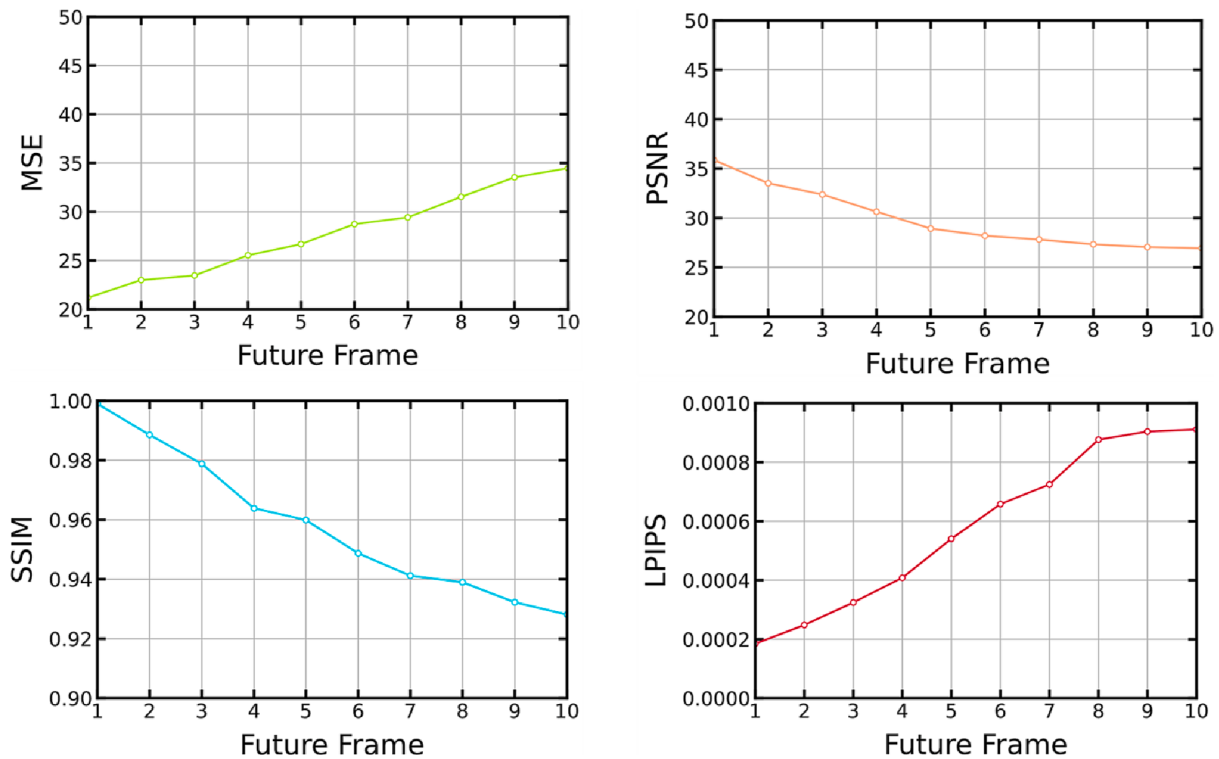
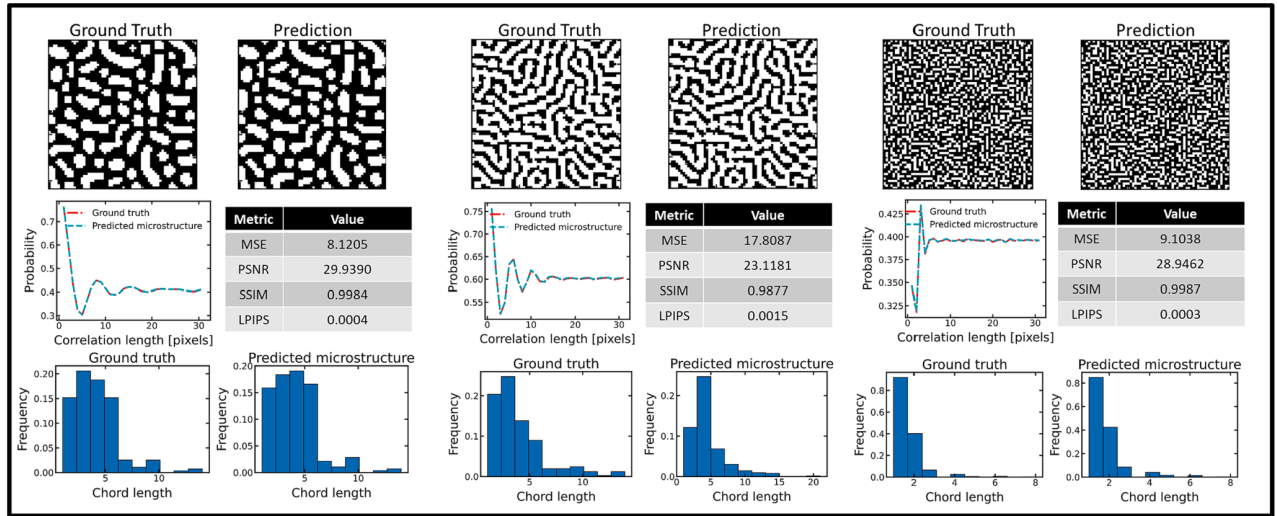


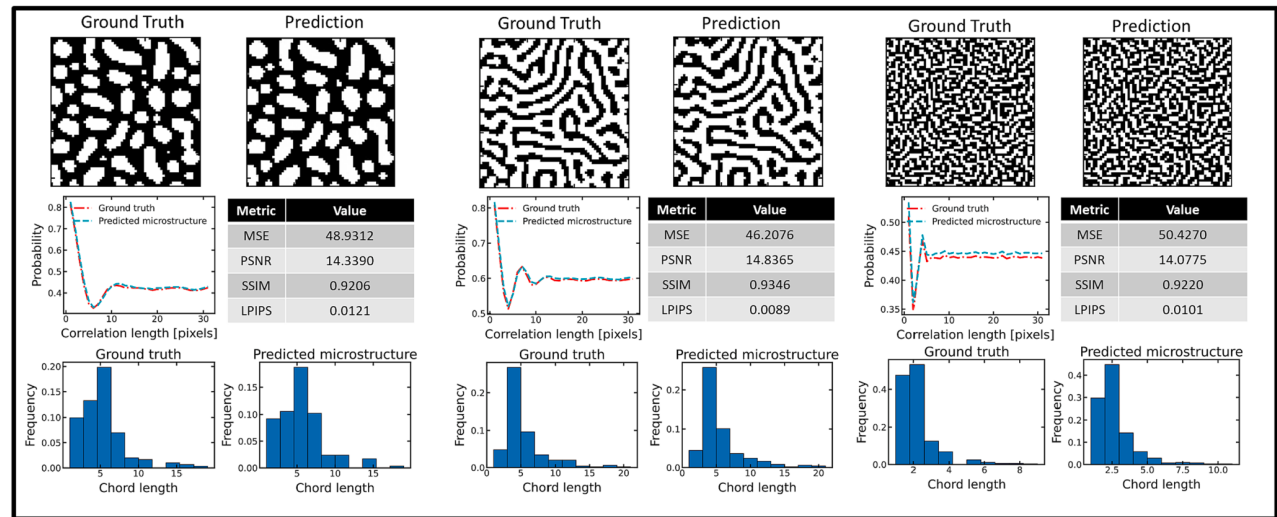
Fig. 7. Frame-wise similarity between the ground truth and PredRNN predicted microstructures on the test set quantified with MSE, PSNR, SSIM, and LPIPS criteria.

trained model. The result for randomly selected test data is given in Fig. 9. While the simulation of the rest microstructures takes more than 75 hrs by the PF modeling, the trained model can predict the future sequence in a fraction of a second by just having earlier microstructures. The error metrics indicate that this prediction is robust and reliable

compared to the simulated microstructures. We note that the developed model in this work predicts the morphology and not the chemistry of the phases, unlike the PF method. However, in our previous works[34,35] we have developed machine learning codes that can predict the chemistry from the morphology image. Therefore, a combination of these two



(a) Short-term prediction



(b) Long-term prediction

Fig. 8. Trained PredRNN model performance on (a) short and (b) long-term prediction for three randomly selected samples from the test set.

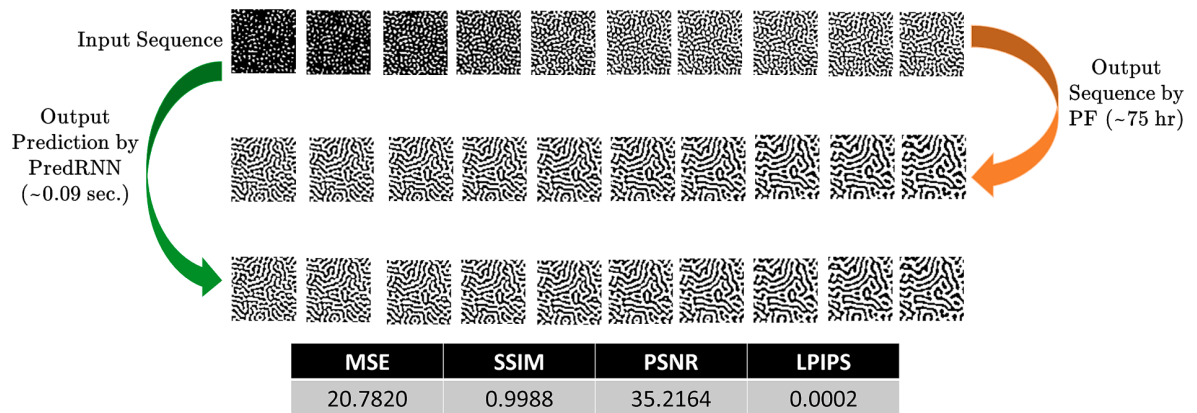


Fig. 9. Comparison of the trained PredRNN model speed with the PF simulation on a randomly selected sample from the test set.

models can predict both morphology and chemistry.

4. Conclusion

We introduced a framework based on a deep neural network to predict the material microstructure evolution. As a case study, we generated the training and testing dataset from phase-field modeling of the spinodal decomposition process in Fe-Cr-Co alloy. We considered the microstructure morphologies evolution based on Fe composition. The future microstructure sequences were predicted by knowing the earlier sequence by PredRNN. Some immediate advantages of the developed framework in on microstructure high throughput simulations and microstructure morphology database generation. The model can predict up to 10X ahead of what it is fed into. This becomes much more critical when we are dealing with multicomponent alloys where the PF simulations get dramatically slow. A detailed analysis of the model's performance indicated that the model parameters were optimized based on training loss reduction and error metrics improvement. The quantitative and qualitative comparisons show that the trained PredRNN model can predict the output sequence accurately. Although the model accuracy for short-term predictions is better than long-term predictions, the model still shows reliable performance in long-term forecasting. The model inference test demonstrates that it can predict the microstructure evolution quickly and accurately. In general, the proposed models could reasonably predict the materials' microstructure evolution.

CRediT authorship contribution statement

Amir Abbas Kazemzadeh Farizhandi: Conceptualization, Methodology, Software, Writing – original draft. **Mahmood Mamivand:** Conceptualization, Investigation, Data curation, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Mahmood Mamivand reports financial support was provided by National Science Foundation.

Data availability

I have shared the link of code and data for this study.

Acknowledgment

The authors appreciate the support of Boise State University. This work was supported in part by the National Science Foundation grant DMR-2142935. We also would like to acknowledge the high-performance computing support of the R2 compute cluster (DOI: 10.18122/B2S41H) that Boise State University's Research Computing Department provided. This work also used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation grant number ACI-1548562.

Data availability

The trained model parameters and dataset to reproduce these findings are available at <https://doi.org/10.24435/materialscloud:es-a4>.

References

- [1] N. Moelans, B. Blanpain, P. Wollants, An introduction to phase-field modeling of microstructure evolution, *Calphad* 32 (2) (2008) 268–294.
- [2] E. Miyoshi, et al., Large-scale phase-field simulation of three-dimensional isotropic grain growth in polycrystalline thin films, *Model. Simul. Mater. Sci. Eng.* 27 (5) (2019), 054003.
- [3] Y. Zhao, et al., Phase-field simulation for the evolution of solid/liquid interface front in directional solidification process, *J. Mater. Sci. Technol.* 35 (6) (2019) 1044–1052.
- [4] J.A. Stewart, R. Dingreville, Microstructure morphology and concentration modulation of nanocomposite thin-films during simulated physical vapor deposition, *Acta Mater.* 188 (2020) 181–191.
- [5] E. Miyoshi, et al., Ultra-large-scale phase-field simulation study of ideal grain growth, *npj Comput. Mater.* 3 (1) (2017) 25.
- [6] Q. Du, X.H. Feng, The phase field method for geometric moving interfaces and their numerical approximations, *Geometric Partial Differential Equations - Part I* (2020).
- [7] E. Herman, J.A. Stewart, R. Dingreville, A data-driven surrogate model to rapidly predict microstructure morphology during physical vapor deposition, *App. Math. Model.* 88 (2020) 589–603.
- [8] X. Zhang, K. Garikipati, Machine learning materials physics: Multi-resolution neural networks learn the free energy and nonlinear elastic response of evolving microstructures, *Comput. Methods Appl. Mech. Eng.* 372 (2020), 113362.
- [9] I. Peivaste, et al., Machine-learning-based surrogate modeling of microstructure evolution using phase-field, *Comput. Mater. Sci.* 214 (2022), 111750.
- [10] D.B. Brough, et al., Extraction of process-structure evolution linkages from X-ray scattering measurements using dimensionality reduction and time series analysis, *Integr. Mater. Manuf. Innovat.* 6 (2) (2017) 147–159.
- [11] G.H. Teichert, K. Garikipati, Machine learning materials physics: Surrogate optimization and multi-fidelity algorithms predict precipitate morphology in an alternative to phase field dynamics, *Comput. Methods Appl. Mech. Eng.* 344 (2019) 666–693.
- [12] Y.C. Yabansu, et al., Application of Gaussian process regression models for capturing the evolution of microstructure statistics in aging of nickel-based superalloys, *Acta Mater.* 178 (2019) 45–58.
- [13] V. Oommen, et al., Learning two-phase microstructure evolution using neural operators and autoencoder architectures, *npj Comput. Mater.* 8 (2022) 1–13.
- [14] M. de Oca, D. Zapiain, J.A. Stewart, R. Dingreville, Accelerating phase-field-based microstructure evolution predictions via surrogate models trained by machine learning methods, *npj Comput. Mater.* 7 (1) (2021) 3.
- [15] Y. Wang, et al., Predrnn: A recurrent neural network for spatiotemporal predictive learning, *IEEE Trans. Pattern Anal. Mach. Intell.* (2022).
- [16] J. Schmidt, et al., Recent advances and applications of machine learning in solid-state materials science, *npj Comput. Mater.* 5 (1) (2019) 83.
- [17] J.M. Rickman, T. Lookman, S.V. Kalinin, Materials informatics: From the atomic-level to the continuum, *Acta Mater.* 168 (2019) 473–510.
- [18] K. Ryan, J. Lengyel, M. Shatruk, Crystal structure prediction via deep learning, *J. Am. Chem. Soc.* 140 (32) (2018) 10158–10168.
- [19] J. Graser, S.K. Kauwe, T.D. Sparks, Machine learning and energy minimization approaches for crystal structure predictions: a review and new horizons, *Chem. Mater.* 30 (11) (2018) 3601–3612.
- [20] P.V. Balachandran, et al., Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning, *Nat. Commun.* 9 (1) (2018) 1668.
- [21] W. Ye, et al., Deep neural networks for accurate predictions of crystal stability, *Nat. Commun.* 9 (1) (2018) 3800.
- [22] T. Xie, J.C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, *Phys. Rev. Lett.* 120 (14) (2018), 145301.
- [23] O. Isayev, et al., Universal fragment descriptors for predicting properties of inorganic crystals, *Nat. Commun.* 8 (1) (2017) 15679.
- [24] F. Yuan, T. Mueller, Identifying models of dielectric breakdown strength from high-throughput data via genetic programming, *Sci. Rep.* 7 (1) (2017) 17594.
- [25] G.P.P. Pun, et al., Physically informed artificial neural networks for atomistic modeling of materials, *Nat. Commun.* 10 (1) (2019) 2339.
- [26] S.M. Azimi, et al., Advanced steel microstructural classification by deep learning methods, *Sci. Rep.* 8 (1) (2018) 2128.
- [27] T. Stan, Z.T. Thompson, P.W. Voorhees, Optimizing convolutional neural networks to perform semantic segmentation on large materials imaging datasets: X-ray tomography and serial sectioning, *Mater. Charact.* 160 (2020), 110119.
- [28] D.S. Bulgarevich, et al., Pattern recognition with machine learning on optical microscopy images of typical metallurgical microstructures, *Sci. Rep.* 8 (1) (2018) 2078.
- [29] G. Ding, et al., A joint deep learning model to recover information and reduce artifacts in missing-wedge sinograms for electron tomography and beyond, *Sci. Rep.* 9 (1) (2019) 12803.
- [30] Y. Mao, et al., High-voltage charging-induced strain, heterogeneity, and micro-cracks in secondary particles of a nickel-rich layered cathode material, *Adv. Funct. Mater.* 29 (18) (2019) 1900247.
- [31] C. Zheng, et al., Random forest models for accurate identification of coordination environments from X-ray absorption near-edge structure, *Patterns* 1 (2) (2020), 100013.
- [32] E.J. Kautz, et al., A machine learning approach to thermal conductivity modeling: A case study on irradiated uranium-molybdenum nuclear fuels, *Comput. Mater. Sci.* 161 (2019) 107–118.
- [33] J.M. Rickman, et al., Materials informatics for the screening of multi-principal elements and high-entropy alloys, *Nat. Commun.* 10 (1) (2019) 2618.
- [34] A.A.K. Farizhandi, O. Betancourt, M. Mamivand, Deep learning approach for chemistry and processing history prediction from materials microstructure, *Sci. Rep.* 12 (1) (2022) 4552.
- [35] A.A.K. Farizhandi, M. Mamivand, Processing time, temperature, and initial chemical composition prediction from materials microstructure by deep network for multiple inputs and fused data, *Mater. Des.* 219 (2022), 110799.

- [36] B.L. DeCost, T. Francis, E.A. Holm, Exploring the microstructure manifold: Image texture representations applied to ultrahigh carbon steel microstructures, *Acta Mater.* 133 (2017) 30–40.
- [37] M. Gusenbauer, et al., Extracting local nucleation fields in permanent magnets using machine learning, *npj Comput. Mater.* 6 (1) (2020) 89.
- [38] A. Cecen, et al., Material structure-property linkages using three-dimensional convolutional neural networks, *Acta Mater.* 146 (2018) 76–84.
- [39] E. Kautz, et al., An image-driven machine learning approach to kinetic modeling of a discontinuous precipitation reaction, *Mater. Charact.* 166 (2020), 110379.
- [40] Z. Yang, et al., Microstructural Materials Design Via Deep Adversarial Learning Methodology, *J. Mech. Des.* (2018).
- [41] W. Ma, et al., Image-driven discriminative and generative machine learning algorithms for establishing microstructure–processing relationships, *J. Appl. Phys.* 128 (13) (2020), 134901.
- [42] G. Liu, et al., Artificial neural network application to microstructure design of Nb-Si alloy to improve ultimate tensile strength, *Mater. Sci. Eng. A* 707 (2017) 452–458.
- [43] L. Exl, et al., Magnetic microstructure machine learning analysis, *J. Phys.: Mater.* (2018).
- [44] A. Barati Farimani, et al. Deep Learning Phase Segregation, 2018. arXiv: 1803.08993.
- [45] P.J. Werbos, Backpropagation through time: what it does and how to do it, *Proc. IEEE* 78 (10) (1990) 1550–1560.
- [46] K. Cho, et al. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches, in: *SSST@EMNLP*, 2014.
- [47] A. Graves, N. Jaitly, Towards end-to-end speech recognition with recurrent neural networks, in: *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, 2014, JMLR.org: Beijing, China. p. II–1764–II–1772.
- [48] J. Donahue, et al., Long-term recurrent convolutional networks for visual recognition and description, *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (4) (2017) 677–691.
- [49] K. Cho, et al. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation, in: *EMNLP*, 2014.
- [50] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, *Neural Comput.* 9 (8) (1997) 1735–1780.
- [51] X. Shi, et al. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, in: *NIPS*, 2015.
- [52] K. Yang, et al., Self-supervised learning and prediction of microstructure evolution with convolutional recurrent neural networks, *Patterns* 2 (5) (2021), 100243.
- [53] Y. Wang, et al. Eidetic 3D LSTM: A Model for Video Prediction and Beyond, in: *ICLR*, 2019.
- [54] E. Moshkelgosha, M. Mamivand, Concurrent modeling of martensitic transformation and crack growth in polycrystalline shape memory ceramics, *Eng. Fract. Mech.* 241 (2021), 107403.
- [55] C.M. Landis, T.J. Hughes, Phase-field modeling and computation of crack propagation and fracture, 2014, TEXAS UNIV AT AUSTIN.
- [56] H. Mehrer, Grain-boundary diffusion, in: *Diffusion in solids: Fundamentals, methods, materials, diffusion-controlled processes*, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 553–582, 2007.
- [57] D.U. Furrer, Application of phase-field modeling to industrial materials and manufacturing processes, *Curr. Opin. Solid State Mater. Sci.* 15 (3) (2011) 134–140.
- [58] L.-Q. Chen, Phase-Field Models for Microstructure Evolution, *Annu. Rev. Mat. Res.* 32 (1) (2002) 113–140.
- [59] S.M. Allen, J.W. Cahn, A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening, *Acta Metall.* 27 (6) (1979) 1085–1095.
- [60] J.W. Cahn, J.E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy, *J. Chem. Phys.* 28 (2) (1958) 258–267.
- [61] T. Koyama, H. Onodera, Phase-Field simulation of phase decomposition in Fe–Cr–Co alloy under an external magnetic field, *Met. Mater. Int.* 10 (4) (2004) 321–326.
- [62] C.J. Permman, et al., MOOSE: Enabling massively parallel multiphysics simulation, *SoftwareX* 11 (2020), 100430.
- [63] J.A. Cornell, Experiments with mixtures: designs, models, and the analysis of mixture data. Vol. 403. 2011, John Wiley & Sons.
- [64] Department, B.S.s.R.C., R2: Dell HPC Intel E5v4 (High Performance Computing Cluster). 2017, Boise State University Boise, ID.
- [65] K.A. Byrne, Borah: Dell HPC Intel (High Performance Computing Cluster), 2020.
- [66] J. Towns, et al., XSEDE: accelerating scientific discovery, *Comput. Sci. Eng.* 16 (5) (2014) 62–74.
- [67] S. Oprea, et al., A Review on Deep Learning Techniques for Video Prediction, *IEEE Trans. Pattern Anal. Mach. Intell.* 44 (6) (2022) 2806–2826.
- [68] J. Oh, et al., Action-conditional video prediction using deep networks in atari games, *Adv. Neural Inf. Proces. Syst.* 28 (2015).
- [69] T. Xue, et al., Visual dynamics: Probabilistic future frame synthesis via cross convolutional networks, *Adv. Neural Inf. Proces. Syst.* 29 (2016).
- [70] J. Zhang, Y. Zheng, D. Qi, Deep spatio-temporal residual networks for citywide crowd flows prediction, in: *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [71] I. Goodfellow, et al., Generative adversarial networks, *Commun. ACM* 63 (11) (2020) 139–144.
- [72] Y. Wu, et al. Future video synthesis with object motion prediction, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
- [73] S. Gur, S. Benaim, L. Wolf, Hierarchical patch vae-gan: Generating diverse videos from a single sample, *Adv. Neural Inf. Proces. Syst.* 33 (2020) 16761–16772.
- [74] B. Liu, et al. Deep learning in latent space for video prediction and compression, in: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.
- [75] M. Ranzato, et al., *Video (language) modeling: a baseline for generative models of natural videos*. arXiv preprint arXiv:1412.6604, 2014.
- [76] N. Srivastava, E. Mansimov, R. Salakhudinov, Unsupervised learning of video representations using lstms. *International conference on machine learning*, PMLR, 2015.
- [77] I. Sutskever, J. Martens, G.E. Hinton, Generating text with recurrent neural networks, in: *ICML*, 2011.
- [78] R. Villegas, et al., High fidelity video prediction with large stochastic recurrent neural networks, *Advances in Neural Information Processing Systems*, 2019. 32.
- [79] J.-Y. Franceschi, et al., Stochastic latent residual video prediction. *International Conference on Machine Learning*, PMLR, 2020.
- [80] B. Wu, et al. Greedy hierarchical variational autoencoders for large-scale video prediction, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.
- [81] R. Villegas, et al., Decomposing motion and content for natural video sequence prediction. arXiv preprint arXiv:1706.08033, 2017.
- [82] N. Bodla, et al. Hierarchical video prediction using relational layouts for human-object interactions, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.
- [83] P. Zablotskaia, et al., Unsupervised video decomposition using spatio-temporal iterative inference. arXiv preprint arXiv:2006.14727, 2020.
- [84] V.L. Guen, N. Thome, Disentangling physical dynamics from unknown factors for unsupervised video prediction, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
- [85] X. Shi, et al., Convolutional LSTM network: A machine learning approach for precipitation nowcasting, *Adv. Neural Inf. Proces. Syst.* 28 (2015).
- [86] C. Finn, I. Goodfellow, S. Levine, Unsupervised learning for physical interaction through video prediction, *Adv. Neural Inf. Proces. Syst.* 29 (2016).
- [87] X. Shi, et al., Deep learning for precipitation nowcasting: A benchmark and a new model, *Adv. Neural Inf. Proces. Syst.* 30 (2017).
- [88] Y. Wang, et al. Eidetic 3D LSTM: A model for video prediction and beyond, in: *International conference on learning representations*, 2018.
- [89] J. Su, et al., Convolutional tensor-train lstm for spatio-temporal learning, *Adv. Neural Inf. Proces. Syst.* 33 (2020) 13714–13726.
- [90] W. Yu, et al., Efficient and information-preserving future frame prediction and beyond, 2020.
- [91] H. Wu, et al. MotionRNN: A flexible model for video prediction with spacetime-varying motions, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.
- [92] A. Graves, N. Jaitly, Towards end-to-end speech recognition with recurrent neural networks. *International conference on machine learning*, PMLR, 2014.
- [93] C. Schudt, I. Laptev, B. Caputo, Recognizing human actions: a local SVM approach, in: *Proceedings of the 17th International Conference on Pattern Recognition*, 2004. ICPR 2004. 2004. IEEE.
- [94] Laboratory, U.N.-E.O. 2017: UCAR/NCAR - Earth Observing Laboratory.
- [95] M. Kopp, et al., Traffic4cast at NeurIPS 2020 - yet more on the unreasonable effectiveness of gridded geo-spatial processes, in: *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*, E. Hugo Jair and H. Katja, Editors. 2021, PMLR: *Proceedings of Machine Learning Research*. p. 325–343.
- [96] Z. Li, et al., Long-Short Term Spatiotemporal Tensor Prediction for Passenger Flow Profile, *IEEE Rob. Autom. Lett.* 5 (4) (2020) 5010–5017.
- [97] F. Amato, et al., A novel framework for spatio-temporal prediction of environmental data using deep learning, *Sci. Rep.* 10 (1) (2020) 22243.
- [98] O. Fagbohunge, L. Qian, Benchmarking inference performance of deep learning models on analog devices, in: *2021 International Joint Conference on Neural Networks (IJCNN)*. 2021. IEEE.