Self-supervised Learning and Prediction of Microstructure Evolution with Recurrent Neural Networks

Kaiqi Yang,1 Yifan Cao,1 Youtian Zhang,1 Ming Tang,1∗
Daniel Aberg,2 Babak Sadigh,2 Fei Zhou2∗

1Department of Materials Science and NanoEngineering, Rice University, Houston, TX 77005, USA
2Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

∗To whom correspondence should be addressed; E-mail: zhou6@llnl.gov, mt20@rice.edu.

Abstract

Microstructural evolution is a key aspect of understanding and exploiting the structure-property-performance relation of materials. Modeling microstructure evolution usually relies on coarse-grained simulations with evolution principles described by partial differential equations (PDEs). Here we demonstrate that convolutional recurrent neural networks can learn the underlying physical rules and replace PDE-based simulations in the prediction of microstructure phenomena. Neural nets are trained by self-supervised learning with image sequences from simulations of several common processes, including plane wave propagation, grain growth, spinodal decomposition and dendritic crystal growth. The trained networks can accurately predict both short-term local dynamics and long-term statistical properties of microstructures and is capable of extrapolating beyond the training datasets in spatiotemporal domains and configurational and parametric spaces. Such a data-driven approach offers significant advantages over PDE-based simulations in time stepping efficiency and offers a useful alternative especially when the material parameters or governing PDEs are not well determined.
Introduction

Materials microstructures are mesoscale structural features that serve as an indispensable link between the atomistic building blocks and macroscopic properties, leading to direct impact on the processing-structure-property relationship of engineered materials. Tailoring materials properties through controlled microstructure evolution under non-equilibrium conditions during materials processing or service, including ubiquitous phenomena such as solidification, solid-state phase transformations and grain growth, is arguably a cornerstone of modern materials science. The ability to understand and predict microstructure evolution has therefore long been a pivotal goal of computational materials design.

Due to time and length scales well beyond the capability of molecular dynamics, simulations of microstructure evolution often rely on coarse-grained models such as partial differential equations (PDEs) as employed in the phase-field method (1). Nevertheless, this approach faces several significant difficulties. First, while in principle PDEs can be derived from the underlying thermodynamic and kinetic considerations, identifying, parametrizing and validating PDEs in practice require intricate evolution rules and substantial manual effort. For complicated or less studied materials, the evolution rules might be either not fully understood or too complex to be described by tractable PDEs. Second, microstructure simulations employing PDEs remain fairly expensive. In the temporal dimension, strict upper limits on the minimum time step size are dictated by the stability of numerical schemes for nonlinear PDEs. In the spatial dimensions, the small grid size determined by the physical interface width often makes direct numerical simulations so expensive that special treatments such as the diffuse interface were introduced (1).

We propose an alternative machine-learning (ML) method to microstructure evolution modeling. Recent progress in ML and deep neural networks (2) in particular enables a data-driven
approach to solving PDEs in place of traditional numerical methods (3–9). Based on statistical learning with big datasets, ML models can be applied without explicit prior knowledge of the physical mechanisms. With proper training, it is possible for ML algorithms to infer “hidden” parameters from the input microstructure images and identify the correct evolution trajectory. Moreover, ML models allow much larger time stepping to achieve significant speedup in the temporal domain. For example, Raissi and coworkers used a single four-layer neural network (3, 4) to obtain the solutions to the Burger’s equation, which otherwise require 500 Runge-Kutta iterations. Breen et al. tackled the notoriously difficult three-body problem with a ten-layer neural nets, skipping thousands of smaller time steps (9). Similarly, coarser spatial grids may be used, as will be shown later. Although previous studies reveal the power of neural nets in rediscovering and solving different types of differential equations, they are mainly limited to ordinary differential equations and PDEs in 1+1 dimensions (i.e. 1 spatial and 1 temporal dimensions). Deep learning of microstructure evolution, which requires PDEs in 2+1 or 3+1 dimensions, remains a challenging subject.

In this work, we apply the recurrent neural networks (RNN) to predict the spatiotemporal evolution of microstructure represented by two-dimensional (2D) image sequences. RNNs are neural nets designed to predict temporal data sequence with hidden memory units (10, 11). With the development of effective variants such as the long short term memory (LSTM) to address the vanishing gradient problem during backpropagation (12), RNNs have found wide-spread success in natural language processing (13, 14), speech recognition (15) and computer vision (16–18). In recent years, several variants of LSTM combined with convolutional neural nets (CNN) have been proposed for predictive learning of spatiotemporal sequences, including the convolutional LSTM (ConvLSTM) (19), Predictive RNN (PredRNN) (20), PredRNN++ (21) and eidetic 3D LSTM (E3D-LSTM) (22). These models employ CNN for efficient spatial
latent-feature extraction and LSTM for feature time evolution to make full use of features in both spatial and temporal domains. We choose the more recent E3D-LSTM method for this study and use the terms E3D-LSTM and RNN interchangeably hereafter.

We assess RNN’s learning ability and predictive power in the context of four well-known evolution phenomena with increasing level of complexity: plane wave propagation, grain growth, spinodal decomposition and dendritic crystal growth. To facilitate comparison with physics-based models, the training datasets are generated from PDE-based simulations or explicit mathematical functions, whose behavior is well understood. A focus of our study is to examine to what degree RNN can grasp and extract the evolution rules from the microstructure images it sees. To this end, extensive and stringent tests are devised to evaluate how well RNN generalizes and extrapolates the learning within the spatiotemporal domain and configurational and parametric spaces. We find that properly trained RNN is able to extend the predictions up to ten folds of the time spans of the training data with significantly larger time step sizes than used in training PDEs, and to systems of larger dimensions. It can forecast the evolution of systems with underlying material parameters or initial configurations that are significantly different from the training images. In addition to excellent piece-wise comparison between the ground truth and short-term predictions, RNN accurately captures the statistical properties of microstructures in the examples considered in the long-term. The satisfactory performance of RNN in these tests provides compelling evidence that it is capable of “comprehending” the physical principles underlying diverse microstructure evolution phenomena, which explains why it is able to make reliable predictions well beyond the scope of training data. Such extrapolation capability further improves the efficiency of RNN by allowing it to be trained with relatively small data size. Our work illustrates the promise of ML approaches in general as a useful alternative to physics-based simulations of microstructure evolution.

Broadly speaking, the use of ML algorithms is growing very rapidly in materials science
in recent years (23–26). They have seen diverse applications ranging from the discovery of new materials (27–31) to the predictions of materials properties (32–36), the development of accurate and efficient potentials for atomistic simulations (37–40), and microscopic and spectroscopic data analysis and processing (41–53). A large number of these works are devoted to materials microstructure, including microstructure classification (41–45), image segmentation (46, 47), predictions of microstructure-property relations (48, 54–56) and microstructure optimization (57–59) with encouraging results. Datasets in these works are mainly in the form of static microstructure images. This work reveals important critical temporal correlation between images of microstructures in their time evolution trajectory.

Results

We employ numerical simulations to generate sequences of 64×64-pixel images as training datasets for four classical examples of evolution phenomena, i.e. plane wave propagation, grain growth, spinodal decomposition and dendritic crystal growth. With varied complexity, they represent a good combination of testing problems for evaluating the capability of RNN in predicting microstructure evolution.

0.1 Plane wave propagation

Before delving into problems pertinent to real materials, we first test RNN with a simple toy model – plane wave propagation dynamics of a scalar field $c$ explicitly described by the following expression:

$$c(x, y, t) = \frac{1}{2} \sin(k_x x + k_y y + \omega t + \theta_0) \exp(-\beta t) + \frac{1}{2}$$

where $\vec{k}=(k_x, k_y)$ is the wave vector, $\theta_0$ is a random phase and $\beta$ is a decay exponent. We use Eq. 1 to generate image sequences, each of which consists of 200 frames at a time interval of
0.005 between two adjacent frames starting at \( t = 0 \). The parameters in Eq. 1 are randomly chosen for each sequence: 

\[
\frac{2\pi}{|\vec{k}|} \in [0.3, 0.6], \quad \frac{2\pi}{\omega} \in [0.03, 0.06], \quad \frac{2\pi}{\beta} \in [1.5, 6] \quad \text{and} \quad \theta_0 \in [0, 2\pi].
\]

Among the generated sequences, 80 are used for training, 20 for validation and 100 for testing. Each simulation sequence is divided into staggered 20-frame training clips (i.e. frame 1–20, 11–30, etc), each of which represents a training data point. For testing, RNN is used to predict the next 50 frames based on an input of 10 consecutive frames. A total of 1500 tests are performed.

Figure 1a illustrates two representative tests, which visually show little difference between the ground truth and predictions. Figure 1b shows the pixel-wise comparison based on the root-mean-squared-error (RMSE) and structural similarity index measure (SSIM) (60) averaged over all of the 1500 tests. Both RMSE and SSIM vary between 0 and 1 and lower RMSE or higher SSIM scores indicate better agreement between the predictions and ground truth. It can be seen that RNN exhibits high piece-wise accuracy in the short-term within the length of training clips, where RMSE stays below 0.5\% and SSIM above 99\%. In the longer term, both RMSE and SSIM vary with time at a greater rate, but remains below 5\% (or above 93\%) up to 50 output frames. As a more revealing measurement of how well RNN recognizes the wave propagation rules, the parameters in Eq. 1 are extracted from the predicted images and compared with their ground truth values. As shown in Figure 1d, the predicted \(|\vec{k}|\) and \(\omega\) differ from the ground truth by less than 2\% but \(\beta\) shows a larger deviation up to 20\%. A probable reason for the predicted \(\beta\) being less accurate is that \(\beta\) characterizes a slower decaying mode of wave motion and may require longer training sequences to learn precisely its temporal behavior.

Overall, RNN exhibits excellent performance when applied to the simple plane wave propagation problem. Next, we test it against more realistic microstructure evolution problems.
Figure 1: **Application of RNN to predicting plane wave propagation.** a. Examples of output frames predicted by the trained RNN (P) based on 10 input frames in comparison with the ground truth (G). b. RMSE (black) and SSIM (blue) of the predictions averaged over 200 testing cases as a function of the frame index \( j \). c. Relative errors of the wave propagation parameters (\(|k|, \omega\) and \(\beta\)) inferred from the predicted images.
0.2 Grain Growth

Grain growth describes the increase of the average grain size in polycrystals with time to reduce the excess energy associated with grain boundaries. During the process, some grains grow while others shrink and disappear, leading to a persisting drop of the number of grains in the system. The growth or shrinkage rate of a grain in 2D polycrystals is determined by its number of sides $N$ according to the famous von Neumann-Mullins or “N-6” rule (61, 62):

$$\frac{dA}{dt} = M\gamma \frac{\pi}{3}(N - 6)$$

where $A$ is the grain area, $M$ and $\gamma$ are the grain boundary mobility and energy, respectively. Eq. 2 states that any grains with fewer than 6 neighbors will shrink and those with more than 6 sides will grow at a rate proportional to $N - 6$.

We generate the training data by performing isotropic 2D grain growth simulations with a phase-field model (63) (see Methods). Simulations are performed on a $256\times256$ grid with periodic boundary conditions to accommodate a sufficient number of grains. Subsequently, the simulation images are down-sampled to $64\times64$ pixels by averaging. Each simulation employs the same parameters but starts with a different initial configuration constructed by Voronoi tessellation with 100 random seeds. It outputs a 20-frame clip after a relaxation period, which serves to remove the artifacts in the polycrystalline structure. The time interval between two adjacent frames corresponds to 80 PDE time steps. The first frame in a clip contains $\sim$75 grains and the last one has $\sim$45 grains. A total of 2400 clips are prepared for training and 600 for validation during training.

After training, RNN is subject to a set of more challenging extrapolation tests than in the wave propagation problem. First, we apply the trained model to predict longer image sequences with less input information. RNN is required to predict 199 frames based on only one input frame. Theoretically, this is feasible as grain growth obeys the dissipation dynamics described
Figure 2: Application of RNN to predicting grain growth.  

**a.** Examples of RNN output frames (P) based on 1 input frame in comparison with the ground truth (G).  

**b.** RNN prediction of the evolution of an artificial polycrystalline configuration, in which four small 4-sided grains are embedded in larger 8-sided grains.  

**c.** RMSE (black) and SSIM (blue) of the predictions averaged over 1000 cases as a function of the frame index \( j \).  

**d.** Time evolution of the average grain area in 1000 testing cases predicted by RNN vs ground truth.  

**e.** Grain size distribution at \( j = 50 \) and 200 predicted by RNN vs ground truth. Effective grain radius is calculated by \( \sqrt{A/\pi} \).
by PDEs of first order in time (Eq. 8). Here the length of the test sequences is ten times of the training clips, and more significantly, 90% of the output frames (frame index $j = 21 – 200$) depict coarsened polycrystalline states never seen by RNN during training. Figure 2a presents two representative tests, which show that RNN does a very good job in the temporal extrapolation. The predictions and ground truth are difficult to distinguish visually in the short term, e.g. at frame index $j = 30$, but visible local structure difference emerges at the later stage. Figure 2c shows that the average RMSE of 1000 tests rises and stabilizes around 20% while SSIM decreases to $\sim 0.4$ at the 200th frame. Despite the increasing difference, the predicted polycrystalline structures are free of any noticeable artifacts throughout the sequences. We note that the accumulation of the discrepancy between the ground truth and predictions is inevitable in the long term. This is because the grain boundary connectivity bifurcates upon grain disappearance (see examples in Supplementary Figure S1), which leads two initially identical configurations onto divergent evolution pathways. As such, statistical measurement of the similarity between two polycrystalline configurations is more meaningful than pixel-wise comparison, and RNN performs very well in this aspect. As shown in Fig. 2d, the error in the predicted average grain area $\langle A \rangle$ of 1000 testing cases remain below 5% while $\langle A \rangle$ has a five-fold increase. Figure 2e shows that the predictions and ground truth also have very good agreement in the grain size distribution. The Euclidean distance between them is only 0.71% at $j = 50$ and still has a low value of 1.61% at $j = 200$. RNN thus faithfully reproduces the statistical characteristics of polycrystals even after a 10-fold extrapolation in time.

Next, we subject RNN to spatial extrapolation tests by asking it to predict grain growth in a system much larger than the training images. Because of the locality of 3D convolution operations in E3D-LSTM, the evolution rules learned by the model can be easily extended to larger domains without additional training. Supplementary Figure S2 presents the results of the grain growth kinetics on a 256×256 mesh predicted by RNN trained on 64×64-pixel images.
The predictions exhibit similar RMSE and SSIM compared to those for the smaller $64 \times 64$-pixel domain. The spatial extensibility of RNN means that there is no need to retrain the model when applying it to problems of different sizes, which is a very appealing feature for practical applications.

As the third type of extrapolation tests, RNN is applied to predict the evolution of artificial polycrystalline configurations qualitatively different from the training data. Figure 2b showcases such an example, in which the system contains four orderly arranged 4-sided grains embedded within four larger grains. Despite the notable morphological difference from those generated by random Voronoi tessellation, the evolution of the polycrystalline is accurately captured by RNN.

The above tests demonstrate RNN’s capability to generalize and extrapolate its learning in the spatiotemporal and configurational spaces. This is a strong indication that it has grasped the evolution rules, which is further supported by other evidence. Grain growth consists of two elementary processes: the continuous shrinkage or expansion of grains without changing their number of sides $N$, and the discontinuous changes in the grain boundary connectivity when grains switch edges or disappear. The former process is governed by the $N - 6$ rule (Eq. 2) resulting from the curvature-driven boundary movement. In Fig. 3a, we show the average growth rates for grains with different $N$ using data from all the 1000 tests. The predictions very faithfully reproduce the $N$ dependence of the ground truth. On the other hand, Fig. 3b illustrates all of the four possible topological events that could occur the grain boundary network upon grain disappearance or edge switching in a 2D system. The numerical examples in Fig. 3b show that RNN correctly predicts each one of them. Therefore, the satisfactory performance of RNN derives from its faithful learning of the elementary steps of the grain growth process.
Figure 3: **Evidence of RNN capturing the evolution rules of grain growth.** a. RNN accurately predicts the dependence of the grain growth rate $\langle dA(N)/dt \rangle$ on the number of grain sides $N$. $\langle dA(N)/dt \rangle$ is averaged over grains of the same $N$ in all of the testing cases. b. Examples from testing cases show that RNN correctly predicts the four possible topological events when a grain disappears or loses an edge to its neighbors. Red circles highlight where the events occur in the predicted images.
0.3 Spinodal decomposition

As a third example of microstructure evolution phenomena, we train RNN to predict spinodal decomposition, which is the phenomenon of spontaneous phase separation in unstable binary or multi-component systems widely found in alloys and polymer blends (64). Mathematically, the spatiotemporal evolution during spinodal decomposition is described by the Cahn-Hilliard (C-H) equation (65) (Eq. 9 in Methods), which is numerically solved to generate the ground truth data in this work. Compared to grain growth, spinodal decomposition is a more complex evolution phenomenon since it involves not only curvature-driven interface migration but also coupled long-range diffusion of chemical species. The complexity is also reflected by the 4th-order nonlinear C-H equation versus the second-order phase-field PDEs for grain growth.

Spinodal decomposition consists of two distinct stages: a fast composition modulation growth stage, followed by a slower coarsening stage, at which the length scale of the phase separation pattern gradually increases due to the Gibbs-Thomson effect (64). We focus on training RNN to recognize the system evolution in the second coarsening stage. Training and validation data are generated from 480 and 120 simulations, respectively, which employ the same parameters but different initial states. The system starts from a uniform binary mixture with one of three compositions at $c_0 = 0.25, 0.5$ and 0.75, which produce different types of domain morphologies. A random noise of the same amplitude is added to the initial configurations to trigger phase separation. Each simulation produces 100 images, and the system becomes phase separated after 2 or 3 frames. Similar to the wave propagation problem, these frames are divided into staggered 20-frame training clips (i.e. frame 1–20, 11–30, ..., and 81–100). The time interval between 2 adjacent frames corresponds to 370 time steps on average in phase-field simulations, which employ an implicit PDE solver with variable time step size.

We perform temporal extrapolation tests on the trained model in a similar way to the case of grain growth. RNN is asked to output 200 frames, or 10 times of the training clip length,
Figure 4: **Application of RNN to predicting spinodal decomposition.** a. Comparison between predictions (P) and ground truth (G) from two testing cases, in which RNN outputs 200 frames based on 1 input frame of spinodal structure generated from random perturbation to a system of uniform composition. b. RNN prediction of the evolution of an artificial bi-phasic configuration, in which second-phase particles ($c = 1$) of randomly chosen radii are orderly arranged within the primary phase ($c = 0$).
given one input frame that is taken from the 50th frame of a simulation starting from a uniform mixture. 75% of the output frames \( (j = 51 – 200) \) thus fall outside the time span of the training sets. In addition, predictions based on 10 input frames are also tested. The results are presented in Supplementary Figure S3, which show similar performance as those with 1 input frame only, which indicates that the information contained in the initial system configuration is sufficient for RNN to correctly project the evolution trajectory. Figure 4a showcases several examples from a total of 510 tests with 170 each having \( c_0 = 0.25, 0.5 \) or 0.75. The short-term predictions up to \( j \sim 50 \) closely resemble the ground truth, which is quantified by the low RMSE \(< 0.06\) and high SSIM \(> 0.97\) in Figure 5a. While the discrepancy gradually accumulates with time and visible difference appears at the later stage, the long-term predictions are realistic looking and no artifacts can be discerned. In addition, Supplementary Figure S4 shows that RNN well conserves the mass in the system, with the average concentration differing less than 6% from \( c_0 \) after 200 output frames. Morphology-wise, it is difficult to tell by human eyes whether the images are generated by RNN or simulations. Such similarity is corroborated by the statistical analysis of the microstructure. In Figure 5b, we compare the interface curvature distributions in the predicted versus ground truth images of 170 testing cases with \( c_0 = 0.5 \), which have a bicontinuous two-phase morphology. The agreement is very good in both short and long terms, which can be quantified by the Euclidean distance between the two distributions: 0.0028 at frame \( j = 26 \) and 0.014 at \( j = 201 \). On the other hand, systems with \( c_0 = 0.25 \) or 0.75 contain individual particles of the minority phase \( (c = 1 \) or 0 \) dispersed within the majority phase. Figure 5c shows the time dependence of the average particle size \( \langle R \rangle \) for 170 tests with \( c_0 = 0.25 \). The corresponding particle size distributions are presented in Figure 5d. The comparison is again satisfactory. The predicted \( \langle R \rangle \) has a maximal error of 1.89% within the test period, and the Euclidean distance between the predicted and true size distributions is only 0.01 at \( j = 26 \) and 0.034 at \( j = 201 \).
We next perform the spatial extrapolation tests by applying the trained model to a larger 256×256-pixel domain. As shown in Supplementary Figure S5, RNN performs equally well in the extended system with comparable RMSE and SSIM as in the smaller domain. Furthermore, Figure 4b shows an example which tests the ability of RNN in predicting the evolution of configurations “foreign” to the training datasets. The initial configuration in the test is created by placing circular particles of $c = 1$ with random radii on square lattice sites in the matrix of $c = 0$. Though never seeing such a configuration during training, RNN captures its evolution very well.

The impressive extrapolation capability of RNN when applied to spinodal decomposition also implies its understanding of the physical rules of this phenomenon. The coarsening of the spinodal structure is thermodynamically driven by the interface curvature dependence of chemical potentials (i.e. Gibbs-Thomson effect) and kinetically limited by the species diffusion. Figure 5b and d show that RNN grasps the Gibbs-Thomson effect, which causes the fraction of low-curvature interface segments to increase with time, and Figure 5c confirms that the diffusion-controlled coarsening kinetics is captured by the model. Apart from the accurate statistical representation, the examples in Figure 5d and e illustrate that RNN is also capable of predicting subtle local morphological changes. The fate of the particle highlighted by red in Figure 5e is determined by the relative sizes of its neighbor particles, which exchange mass between each other via diffusion due to the size-dependent chemical potential. The red particle first grows at the expense of a smaller neighbor, but subsequently shrinks by losing mass to the other two bigger particles nearby. In Figure 5f, the particle in red receives an incoming diffusion flux from two smaller adjacent particles. Its growth rate exhibits two bursts, which coincide with the complete dissolution of the two particles. RNN’s ability to predict detailed evolution features as demonstrated in these examples further inspires confidence in its comprehension of the underlying physics.
Figure 5: **Accuracy of RNN in predicting spinodal decomposition.** a. RMSE and SSIM of RNN predictions averaged over 510 testing cases as a function of the frame index $j$. b. Distribution of the interface segment curvature $\kappa$ at $j = 26$ and $201$ in 170 testing cases with $c_0 = 0.5$. c. and d. Evolution of the average second-phase particle radius $\langle R \rangle$ (c) and particle area $A_p$ distribution (d) in 170 testing cases with $c_0 = 0.25$. $\langle R \rangle$ is calculated as $\sqrt{\langle A_p \rangle / \pi}$. e. and f. (top) Examples of local morphological evolution predicted by RNN from two testing cases with $c_0 = 0.25$. (bottom) Size evolution of the red particle in the images as predicted by RNN vs ground truth.
0.4 Dendrite growth

In the last example, we give RNN a more challenging task to predict dendritic crystallization patterns. During crystal growth, dendritic structures like the beautiful snowflakes often form due to the morphological instability of the growth front, which is promoted by the negative temperature and/or species concentration gradient(s) ahead of the phase boundary and the interface energy anisotropy. Such instability phenomena are intrinsically difficult to predict. In addition, dendrite growth is a multiphysical process coupling phase transformation, long-range mass and heat transport and interface instability. As a result, microstructure images fed to RNN do not contain the complete information of the system state, which further increases the difficulty of making accurate predictions.

Here we generate training data using a phase-field model of solidification in pure systems by Kobayashi (66). As described in the Methods section, the spatiotemporal evolution of the system state is described by two coupled PDEs for the temperature ($T$) and phase-field ($\phi$) variables. $\phi$ distinguishes between the solid ($\phi=1$) and liquid ($\phi=0$) phases during solidification. We use $\phi(t, x, y)$ to create the microstructure images. $T$ and other parameters in the governing equation (Eq. 15) such as the normalized latent heat $K$ are thus hidden to the learning process. We perform phase-field simulations on a $64 \times 64$ mesh, in which a small solid nucleus is placed at or near the center and surrounded by the supercooled liquid phase. The training and validation sets contain 800 and 200 simulations, respectively. To enrich the training data, each simulation has a different nucleus, crystal orientation $\theta_0$ and $K$. Specifically, $K$ is randomly chosen from (1.2, 2) and $\theta_0$ from (0, $\pi/3$) (crystal is assumed to have six-fold symmetry). The nucleus is given random shape (circle, rectangle or ellipse), size (2 – 6 pixels) and off-center distance ($\pm 5$ pixels in $x$ and $y$ directions). Similar to the case of spinodal decomposition, 100 image frames with equal time interval are obtained from a simulation and divided into eight staggered 20-frame training clips.
Figure 6: Application of RNN to predicting dendritic crystal growth. RNN predictions (P) vs. ground truth (G) from five testing cases with different $K$ values, in which RNN outputs 50 frames based on 10 input frames.
In testing, the trained RNN model is required to predict 50 frames from 10 consecutive input frames, which are taken from the first half of a simulation. Predictions are not extended to longer time because the dendrite tips already approach the domain boundaries after 50 output frames in many tests and growth stagnates subsequently. Instead, we focus on conducting the extrapolation tests in the model parameter space. Specifically, $K$ is randomly and uniformly selected from $(0.8, 2.4)$ to generate ground truth data in the testing cases. This means that half of the selected $K$ values fall outside its range in the training data, which is $(0.8, 2)$. $\theta_0$ and the solid nucleus shape are also randomized. **Figure 6** presents several examples from a total of 600 testing cases. The predicted dendritic structure matches the ground truth well in all the cases even at $K = 1.161$ and 2.106, which are outside the scope of training data. In particular, RNN captures the fine features of the dendrites such as the locations of secondary side branches. It can be seen that the crystal growth pattern depends strongly on $K$. Smaller $K$ results in thicker primary branches and more compact morphology. RNN manages to recognize the correct evolution trajectory based on the input images without prior knowledge of the underlying $K$ value.

**Figure 7a** shows RMSE and SSIM of the predictions averaged over all of the 600 testing cases. RNN fares well in pixel-wise comparisons although the prediction error increases faster with time than in the cases of grain growth and spinodal decomposition, which can be attributed to the more complex physics of the dendrite growth process.

As a more revealing indicator of RNN’s performance, we use several shape descriptors (Feret diameter $d_{\text{feret}}$, convexity, solidity) to characterize the dendrite morphology. Feret diameter, which is defined as the maximum distance between two parallel tangent lines touching the shape, provides a measure of the linear dendrite dimension. Convexity and solidity quantify the degrees of concavity and compactness of the crystal. **Figure 7b** shows the time evolution of these descriptors from one test while their average errors for all the 600 tests are plotted in **Figure 7c**. It can be seen that RNN accurately predicts the dendritic shape evolution with the
Figure 7: **Accuracy of RNN in predicting dendritic crystal growth.** a. RMSE (black) and SSIM (blue) of the predictions averaged over 600 testing cases. b. Time evolution of the Feret diameter $d_{\text{feret}}$, convexity and solidity of a growing crystal from a testing case. Solid lines are ground truth (G), and dashed lines are predictions (P). The shape descriptors are calculated in imageJ after image binarization. Convexity is defined as $L_h/L_c$, where $L_c$ is the crystal perimeter and $L_h$ is the perimeter of the convex hull of the crystal. Solidity is defined as $A_c/A_h$, where $A_c$ is the crystal area and $A_h$ is the area surrounded by the convex hull of the crystal shape. c. Relative errors of predicted $d_{\text{feret}}$, convexity and solidity of crystals averaged over 600 testing cases as a function of image frame index $j$. d. Development of secondary branches on the dendritic crystal in a testing case. $N_s$ is the number of secondary branches on a primary branch of the dendrite. Insets above and below the curves show the crystal shape from the ground truth and predictions, respectively, at times marked by the black squares.
average error less than 7% throughout the tests. In addition to global metrics, we also examine how well RNN reproduces local dendritic structural features. In Figure 7d, the number of secondary branches formed on a primary branch in a test is plotted as a function of time. It shows that RNN performs very well in predicting the frequency of the side-branching events occurring near the dendrite tip.

**Discussion**

In addition to the prediction accuracy, we compare the computational efficiency of using RNN for microstructure evolution predictions with that of PDE-based simulations. During training and testing, the time interval between two RNN output frames is 80 times that of the average time step size used in the grain growth simulation, 370 times in spinodal decomposition and 7 times in dendritic crystal growth. These large time spacings will quickly render PDE solvers unstable. This illustrates a significant computational benefit of RNN, that its performance is not limited by the numerical stability of PDEs and can thus make reliable predictions at much larger time step size. In the grain growth example, RNN’s advantage in spatial coarsening is also demonstrated. Because of the diffuse-interface representation, a grain boundary needs to be resolved by at least 5–6 pixels in phase-field simulations to maintain desired numerical accuracy (67). However, RNN is not subject to the same spatial resolution requirement and can predict system evolution on a coarser mesh (64×64) than used in phase-field simulations (256×256). Accordingly, the improved efficiency in time marching and spatial coarse-graining results in computational saving. We benchmark E3D-LSTM’s running time on a compute node with 4 NVidia GeForce GTX 1080-TI GPUs. After spending 130 – 450 seconds to initialize and load pre-trained models, it takes on average 2.1 s to predict a case for grain growth (200 64×64-pixel output frames), 3.8 s for spinodal decomposition (200 64×64-pixel output frames), and 0.56 s for dendrite growth (50 64×64-pixel output frames). In comparison, numerical
simulations of the three examples, which are implemented by an in-house C code (grain growth) and COMSOL Multiphysics 5.3 (spinodal decomposition, dendrite growth) on a desktop with an Intel i7 3.2GHz CPU, require an average running time of 1551 s for grain growth, 350 s for spinodal decomposition and 44 s for dendrite growth. The comparison shows that RNN is computationally efficient, especially when applied to a large number of cases so that the overhead associated with initialization is small. Note that the speedup estimation made here is conservative: the PDEs employed were simple textbook examples and substantially cheaper than more sophisticated equations in real applications, and were implemented in efficient C codes, while the neural nets were based on research Python codes and not optimized for physics simulation.

The overall efficiency of RNN in predicting microstructure evolution also depends on the training data size and the efforts and resources required for data collection. Supplementary Figure S6 shows the dependence of validation error on the number of training clips $N_{\text{clip}}$ in the cases of plane wave propagation and grain growth. In both cases, the improvement in model performance becomes negligible after $N_{\text{clip}}$ goes beyond $\sim 2000$. On the other hand, we find that increasing the length of training clips beyond 20 frames does not significantly improve the prediction accuracy. For all of the examples in this work, the time spent on generating the training datasets is comparable to the model training time. Therefore, the data requirement of RNN should not present a major obstacle to its applications.

Despite the overall very impressive performance, our tests show that the learning rate and predictive power of RNN vary with the nature of the microstructure evolution phenomena it is applied to. Among all the examples, RNN demonstrates the best learning ability in predicting grain growth because its evolution rules are localized, which could be relatively easily recognized by E3D-LSTM through 3D convolution operations that specialize in remembering local, short-term motion. In contrast, training RNN to predict spinodal decomposition is more
challenging because the long-range mass transport inherent in the process creates longer and stronger spatiotemporal correlation, which requires more convolution operations and long-term memory states to extract the essential features. In fact, the model can be successfully trained to predict grain growth with only two E3D-LSTM layers, but 4 layers are needed for spinodal decomposition to reach similar performance. Compared to grain growth, we also find it necessary to include longer image sequences (100 frames) into the training datasets for spinodal decomposition to better inform RNN of the evolution trajectories and achieve comparable prediction accuracy. Predicting dendrite growth presents additional challenges due to the interface instability and the existence of hidden variable \( T \) not directly seen by RNN. However, potential improvement could be achieved by encoding both \( \phi(t, x, y) \) and \( T(t, x, y) \) into multi-channel images to let RNN learn the evolution of not only the microstructure morphology but also other relevant fields, which could be a general strategy to effectively predict microstructure evolution governed by complex multiphysics principles.

In summary, we train a convolutional recurrent neural network (E3D-LSTM) to predict the spatiotemporal evolution of materials microstructure. Using training data from four distinct evolution processes (plane wave propagation, grain growth, spinodal decomposition and dendritic crystal growth), the RNN, which is composed of the same network architecture and agnostic about the underlying physics, was able to adapt efficiently to different evolution rules. The ability of RNN to generalize learning beyond the training datasets is systematically examined by a series of extrapolation tests, even though the neural net is , In addition to performing very well in piecewise comparison with ground truth in short-term predictions, RNN accurately describes the statistical properties of microstructures over long periods up to ten folds of the training data’s time span. Without additional training, neural nets trained on small-size images can be straightforwardly applied to larger systems with comparable accuracy. The method can reliably predict the evolution of microstructures whose morphology or underlying materials parameters
differ qualitatively from the training data. The spatiotemporal, configurational and parametric extensibility demonstrated by RNN suggests that it is capable of learning the evolution rules of the microstructure phenomena considered here, which provides the physical basis for its practical applications. Computationally, RNN is not restricted by the numerical stability of PDE solvers and can employ time step size 1-2 orders of magnitude larger than PDE-based simulations in our tests. The ML approach demonstrated in this study provides a valuable complement and/or alternative to physics-based simulations for predicting microstructure evolution, which could be especially attractive in situations where there exist unknown materials parameters or evolution principles are not fully understood.

**Methods**

**I. Recurrent Neural Network**

Unlike static data without temporal context, sequential data such as the microstructure evolution trajectories in the form of image sequences require special treatment for deep neural networks to learn efficiently and accurately. Designed to take advantage of the temporal information of sequential inputs, RNN along with its LSTM variants were first successfully employed in voice recognition and natural language processing. Recently, Shi et al. (19) proposed a convolutional LSTM model for image sequence prediction, which uses CNN instead of fully connected layers in vanilla RNN for latent-feature extraction, and combines it with LSTM for learning time evolution to make full use of features in both spatial and temporal domains. More recent studies replace the initially stacked chain structure (19) with sophisticated neural nets to improve information flow and reach better performance.

For example, Yunbo Wang and co-workers developed a series of neural networks for spatiotemporal predictive learning (20–22). The latest Eidetic 3D LSTM (E3D-LSTM) model is employed in our study. Compared with other state-of-the-art models that use 2D convolution
operations, E3D-LSTM integrates 3D (one temporal and two spatial dimensions) convolution (3D-Conv) deep into RNNs, which is effective for modeling local representations in a consecutive manner. As shown in Figure 1(c) of Ref. (22), successive input frames are encoded by 3D-Conv encoders before being fed to E3D-LSTM units. Outputs of E3D-LSTM units are decoded with a 3D-Conv layer to obtain the real-space predictions. Besides adopting 3D-Conv as its basic operations, E3D-LSTM exploits a self-attention mechanism to memorize long-term interactions in addition to short-term motions. This is achieved by implementing two distinct memory states in E3D-LSTM: spatiotemporal memory and eidetic 3D memory. The former is designed to capture the short-term motion (20) while the latter computes the relation between local patterns and the whole memory space to distinguish and revoke temporally distant memories.

**Model setup:** Each data point in the training sets is a sequence of $N_t$ 2D images generated by a scalar field $c(t, x, y)$ ($0 \leq c \leq 1$, $t = 1 \ldots N_t$, $x = 1 \ldots N_x$, $y = 1 \ldots N_y$). The spatial dimensions $N_x$ and $N_y$ are 64 unless otherwise stated. For each problem considered, the training dataset is a 4D array $c_i(t, x, y)$ with $N_{total}$ image sequences ($i = 1 \ldots N_{total}$). Following Ref. (22), four E3D-LSTM layers are stacked together in the model (only two layers in the case of grain growth), each with 64 hidden features. The model is implemented in Tensor Flow (68) and trained on 4 NVidia V100 or 1080-TI GPUs. Typical training time is 36–48 hours, with an initial learning rate of $10^{-3}$ that gradually decays to $10^{-5}$.

**Data augmentation:** Training data are augmented by performing symmetry operations of the 2D point group $4mm$ on the original images, which transform $c(t, x, y)$ to $c(t, \bar{x}, y)$, $c(t, x, \bar{y})$, $c(t, \bar{x}, \bar{y})$, $c(t, y, x)$, $c(t, \bar{y}, x)$, $c(t, y, \bar{x})$ and $c(t, \bar{y}, \bar{x})$ ($\bar{x} \equiv N_x + 1 - x$, $\bar{y} \equiv N_y + 1 - y$). Such data transformations can be achieved by array rearrangements and do not require additional float-point calculations.

**Analysis methods:** RMSE and SSIM are used in pixel-wise comparison between ground
truth and predictions. RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (p_g(i,j) - p_p(i,j))^2}{N_xN_y}}$$

(3)

where $p_g(i,j)$ and $p_p(i,j)$ are the pixel values of ground truth and predictions, respectively.

SSIM (60) is defined as

$$\text{SSIM} = \frac{(2\bar{p}_g\bar{p}_p + c_1)(2\sigma_{gp} + c_2)}{\overline{(\sigma_g^2 + \sigma_p^2 + c_1)}(\sigma_g^2 + \sigma_p^2 + c_2)}$$

(4)

where $\bar{p}_k$, $\sigma_k$ ($k = g, p$) are the average pixel value and variance of ground truth or predictions, respectively, and $\sigma_{gp}$ is their covariance. $c_1$ and $c_2$ are small constants and chosen to be $c_1 = (0.01L)^2$ and $c_2 = (0.03L)^2$, where $L$ is the range of pixel values. The Euclidean distance between the distributions of quantity $q$ from RNN predictions and ground truth is defined as

$$d = \sqrt{\sum_{i=1}^{n} (q^i_g - q^i_p)^2}$$

(5)

where $n$ is the number of bins within the interval between the minimum and maximum of $q$, and $q^i_g$ and $q^i_p$ are normalized counts in the $i$-th bin of the ground truth and predictions, respectively. $n = 20$ is used for all the calculations.

II. Simulation method

Phase-field simulations are employed to generate the ground truth for three microstructure evolution processes, i.e. grain growth, spinodal decomposition and dendritic crystal growth. Phase-field method is a powerful computational technique for modeling microstructure evolution in diverse materials systems (67, 69, 70). In a phase-field model, different phases are represented by one or multiple order parameters, and their interfaces are tracked by the level sets of the order parameters. Spatiotemporal evolution of the microstructure is described by the governing equations of the order parameters derived from thermodynamic and kinetic principles.
Grain growth: Isotropic grain growth in 2D polycrystalline structure is simulated by a multi-order-parameter phase-field model (63). In the model, a set of order parameters \( \{ \eta_1(x), \eta_2(x), ..., \eta_N(x) \} \) are used to represent \( N \) distinct grain orientations. The free energy of the system is expressed as

\[
F = \int \left[ f(\eta_1, \eta_2, ..., \eta_N) + \nu \sum_{i=1}^{N} (\nabla \eta_i)^2 \right] dV
\]

(6)

where the homogeneous free energy density \( f \) is given by

\[
f = m \left[ \sum_{i=1}^{N} \left( \frac{\eta_i^4}{4} - \frac{\eta_i^2}{2} \right) + \frac{3}{2} \sum_{i=1}^{N} \sum_{j>i}^{N} \eta_i^2 \eta_j^2 + \frac{1}{4} \right]
\]

(7)

which has \( N \) local minima located at \((\eta_1, \eta_2, ..., \eta_N) = (1, 0, ..., 0), (0, 1, ..., 0), ..., (0, 0, ..., 1)\). The evolution of \( \eta_i(x) \) \((i = 1 \ldots N)\) follows the time-dependent Ginzburg-Landau or Allen-Cahn (71, 72) equation

\[
\frac{\partial \eta_i}{\partial t} = -L \frac{\delta F}{\delta \eta_i}
\]

(8)

In all the simulations, dimensionless parameters \( N = 100, m = 1, \nu = 1 \) and \( L = 1 \) are used. The initial polycrystalline structure is generated by Voronoi tessellation (73) with 100 grains. Eq. 8 is solved by the forward Euler finite difference scheme with periodic boundary conditions and grid spacing \( \Delta x = 1 \) and time step size \( \Delta t = 0.2 \). Single-channel images of the polycrystalline structure are generated by assigning \( \sum_{i=1}^{N} \eta_i^3 \) as the pixel value so that pixels are close to 0 in the grain boundary region and 1 inside grains.

Spinodal decomposition: Spinodal decomposition is simulated by the Cahn-Hilliard equation (65),

\[
\frac{\partial c}{\partial t} = \nabla \cdot \left[ Mc(1-c) \nabla \left( \frac{\partial f_{chem}}{\partial c} - \epsilon \nabla^2 c \right) \right],
\]

(9)

where \( c \) is the molar fraction of a species in a binary system. We use the regular solution model to describe the homogeneous free energy density:

\[
f_{chem}(c) = RT[c \ln c + (1-c) \ln(1-c)] + \omega c(1-c)
\]

(10)
with a positive value is assigned to the regular solution coefficient $\omega$ to favor phase separation. Eqs. (9)-(10) are solved with no-flux boundary conditions. Dimensionless parameter values $\omega = 0.27397$, $\epsilon = 0.1682$ and $M = 1$ and mesh spacing $\Delta x = 1$ are used in all of the simulations. Eq. 9 is solved with an implicit backward differentiation formula (BDF) solver in COMSOL Multiphysics with an average dimensionless time step size of 4.01. Images are output from simulations at a time interval of 1500, or an average of 370 steps between two frames.

**Dendrite growth:** We use a phase-field model developed by Kobayashi (66) to simulate the dendritic solidification process in a pure material system. Compared to other more quantitative models (70, 74), this model is chosen for its simplicity since the purpose of this work is not to study dendritic growth but use it as an example to evaluate RNN. The system state is described by the temperature field $T$ and an order parameter $\phi$, which distinguishes between the solid ($\phi = 1$) and liquid ($\phi = 0$) phases. The free energy of the system is given by

$$F[\phi, T] = \int \left[ \frac{1}{2} \epsilon(\theta)^2 |\nabla \phi|^2 + f(\phi, T) \right] \, dx,$$

(11)

where the anisotropy of the solid/liquid interface energy is controlled by the orientation dependence of the gradient energy coefficient: $\epsilon(\theta) = \epsilon_0 (1 + \delta \cos[n(\theta - \theta_0)])$, where $\theta$ represents the interface normal and is calculated from the gradient of $\phi$ as $\theta = \arctan(-\phi_y/\phi_x)$. We employ $n = 6$ in simulations to produce dendrites with sixfold symmetry. $f$ is a double-well potential

$$f(\phi, T) = \frac{1}{4} \phi^4 - \left[ \frac{1}{2} - \frac{1}{3} m(T) \right] \phi^3 + \left[ \frac{1}{4} - \frac{1}{2} m(T) \right] \phi^2,$$

(12)

$$m(T) = \frac{\alpha}{\pi} \arctan[\gamma(T_{eq} - T)].$$

(13)

where $T_{eq}$ is the solid/liquid equilibrium temperature. The time evolution of the coupled $\phi$ and
$T$ fields is governed by

$$\tau \frac{\partial \phi}{\partial t} = -\frac{\delta F}{\delta \phi},$$

(14)

$$\frac{\partial T}{\partial t} = \nabla^2 T + K \frac{\partial \phi}{\partial t}$$

(15)

where constant $K$ represents the latent heat. The following dimensionless parameters are used in all the simulations: $\alpha = 0.9$, $\gamma = 10$, $T_{eq} = 1$, $\tau = 0.001$, $\epsilon_0 = 0.01$, $\delta = 0.03$ while $K$ and $\theta_0$ are varied. The system has a uniform initial temperature at $T(t = 0, x, y) = 0$. Eq. 14 and 15 are solved with a BDF solver in COMSOL Multiphysics with mesh spacing $\Delta x = 1$ and average time step size $\Delta t = 5.7 \times 10^{-4}$. Images are output from simulations at a time interval of 0.004, or an average of 7 time steps between two frames.

**Data availability**

The data that support the findings of this study are available from the corresponding authors upon request.

**Acknowledgements**

K.Y., Y.C and M.T. acknowledge support from DOE under project number DE-SC0019111. Y.Z. acknowledges support from NSF under project number CMMI-1929949. The work of D.A., B.S. and F.Z. was supported by the Critical Materials Institute, an Energy Innovation Hub funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Advanced Manufacturing Office, and performed under the auspices of the U.S. Department of Energy by LLNL under Contract DE-AC52-07NA27344. Phase-field simulations were performed on supercomputers at the Texas Advanced Computing Center (TACC) at The University of Texas. RNN training and testing were performed on supercomputers at LLNL and TACC.
Author contributions

F.Z. conceived the project. F.Z. and M.T. supervised the project. K.Y., Y.Z. and Y.C. performed phase-field simulations. K.Y., Y.C. and F.Z. performed RNN training and testing. K.Y., Y.C., Y.Z., M.T., D.A., B.S. and F.Z. analyzed and discussed the results. M.T., F.Z., K.Y. and Y.C. wrote the manuscript with inputs from other authors.

References

1. N. Provatas, K. Elder, Phase-Field Methods in Materials Science and Engineering (John Wiley & Sons, Ltd, 2010).

2. Y. LeCun, Y. Bengio, G. Hinton, Nature 521, 436 (2015).

3. M. Raissi, P. Perdikaris, G. E. Karniadakis, arXiv.org (2017).

4. M. Raissi, P. Perdikaris, G. E. Karniadakis, arXiv.org (2017).

5. S. L. Brunton, J. L. Proctor, J. N. Kutz, Proc. Natl. Acad. Sci. 113, 3932 (2016).

6. S. H. Rudy, S. L. Brunton, J. L. Proctor, J. N. Kutz, Science Advances 3, e1602614 (2017).

7. K. Champion, B. Lusch, J. N. Kutz, S. L. Brunton, Proc. Natl. Acad. Sci. 116, 22445 (2019).

8. H. Schaeffer, Proc. R. Soc. A 473, 20160446 (2017).

9. P. G. Breen, C. N. Foley, T. Boekholt, S. P. Zwart, arXiv.org (2019).

10. D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning Representations by Back-Propagating Errors (MIT Press, Cambridge, MA, USA, 1988), p. 696699.

11. P. J. Werbos, Proceedings of the IEEE 78, 1550 (1990).
12. S. Hochreiter, J. Schmidhuber, *Neural computation* 9, 1735 (1997).

13. I. Sutskever, J. Martens, G. E. Hinton, *ICML* (2011).

14. K. Cho, B. Van Merriënboer, D. Bahdanau, Y. Bengio, *arXiv preprint arXiv:1409.1259* (2014).

15. A. Graves, N. Jaitly, *International conference on machine learning* (2014), pp. 1764–1772.

16. N. Srivastava, E. Mansimov, R. Salakhudinov, *International conference on machine learning* (2015), pp. 843–852.

17. J. Donahue, *et al.*, *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 2625–2634.

18. C. Finn, I. Goodfellow, S. Levine, *Advances in neural information processing systems* (2016), pp. 64–72.

19. S. Xingjian, *et al.*, *Advances in neural information processing systems* (2015), pp. 802–810.

20. Y. Wang, M. Long, J. Wang, Z. Gao, P. S. Yu, *NIPS* (2017).

21. Y. Wang, Z. Gao, M. Long, J. Wang, P. S. Yu, *ICML* (2018).

22. Y. Wang, *et al.*, *ICLR* (2019).

23. J. Schmidt, M. R. Marques, S. Botti, M. A. Marques, *npj Computational Materials* 5, 1 (2019).

24. J. M. Rickman, T. Lookman, S. V. Kalinin, *Acta Materialia* 168, 473 (2019).

25. R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, *npj Computational Materials* 3, 1 (2017).
26. K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, A. Walsh, *Nature* **559**, 547 (2018).

27. K. Ryan, J. Lengyel, M. Shatruk, *Journal of the American Chemical Society* **140**, 10158 (2018).

28. J. Graser, S. K. Kauwe, T. D. Sparks, *Chemistry of Materials* **30**, 3601 (2018).

29. P. V. Balachandran, B. Kowalski, A. Sehirlioglu, T. Lookman, *Nature communications* **9**, 1 (2018).

30. W. Ye, C. Chen, Z. Wang, I.-H. Chu, S. P. Ong, *Nature communications* **9**, 1 (2018).

31. F. A. Faber, A. Lindmaa, O. A. Von Lilienfeld, R. Armiento, *Physical review letters* **117**, 135502 (2016).

32. T. Xie, J. C. Grossman, *Physical review letters* **120**, 145301 (2018).

33. O. Isayev, et al., *Nature communications* **8**, 1 (2017).

34. F. Yuan, T. Mueller, *Scientific reports* **7**, 1 (2017).

35. C. Kim, G. Pilania, R. Ramprasad, *The Journal of Physical Chemistry C* **120**, 14575 (2016).

36. J. Carrete, W. Li, N. Mingo, S. Wang, S. Curtarolo, *Physical Review X* **4**, 011019 (2014).

37. G. P. Pun, R. Batra, R. Ramprasad, Y. Mishin, *Nature communications* **10**, 1 (2019).

38. J. Behler, *The Journal of chemical physics* **145**, 170901 (2016).

39. V. Botu, R. Ramprasad, *Physical Review B* **92**, 094306 (2015).

40. J. Behler, M. Parrinello, *Physical review letters* **98**, 146401 (2007).

41. B. L. DeCost, T. Francis, E. A. Holm, *Acta Materialia* **133**, 30 (2017).
42. B. L. DeCost, E. A. Holm, *Computational materials science* **110**, 126 (2015).

43. A. Chowdhury, E. Kautz, B. Yener, D. Lewis, *Computational Materials Science* **123**, 176 (2016).

44. J. Ling, *et al.*, *Materials Discovery* **10**, 19 (2017).

45. S. M. Azimi, D. Britz, M. Engstler, M. Fritz, F. Mücklich, *Scientific reports* **8**, 1 (2018).

46. T. Stan, Z. T. Thompson, P. W. Voorhees, *Materials Characterization* **160**, 110119 (2020).

47. D. S. Bulgarevich, S. Tsukamoto, T. Kasuya, M. Demura, M. Watanabe, *Scientific reports* **8**, 1 (2018).

48. J. M. Rickman, Y. Wang, A. D. Rollett, M. P. Harmer, C. Compson, *npj Computational Materials* **3**, 1 (2017).

49. M. Ziatdinov, A. Maksov, S. V. Kalinin, *npj Computational Materials* **3**, 1 (2017).

50. M. Ziatdinov, *et al.*, *ACS nano* **11**, 12742 (2017).

51. G. Ding, Y. Liu, R. Zhang, H. L. Xin, *Scientific reports* **9**, 1 (2019).

52. Y. Mao, *et al.*, *Advanced Functional Materials* **29**, 1900247 (2019).

53. C. Zheng, C. Chen, Y. Chen, S. P. Ong, *Patterns* p. 100013 (2020).

54. M. Gusenbauer, *et al.*, *npj Computational Materials* **6**, 1 (2020).

55. R. Kondo, S. Yamakawa, Y. Masuoka, S. Tajima, R. Asahi, *Acta Materialia* **141**, 29 (2017).

56. A. Cecen, H. Dai, Y. C. Yabansu, S. R. Kalidindi, L. Song, *Acta Materialia* **146**, 76 (2018).

57. R. Liu, *et al.*, *Scientific reports* **5**, 1 (2015).
58. G. Liu, et al., *Materials Science and Engineering: A* **707**, 452 (2017).

59. L. Exl, et al., *Journal of Physics: Materials* **2**, 014001 (2018).

60. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, et al., *IEEE transactions on image processing* **13**, 600 (2004).

61. J. Von Neumann, *American Society for Metals, Cleveland* **108** (1952).

62. W. W. Mullins, *Journal of Applied Physics* **27**, 900 (1956).

63. N. Moelans, B. Blanpain, P. Wollants, *Physical Review B* **78**, 024113 (2008).

64. R. W. Balluffi, S. M. Allen, W. C. Carter, *Kinetics of materials* (John Wiley & Sons, 2005).

65. J. W. Cahn, J. E. Hilliard, *The Journal of chemical physics* **28**, 258 (1958).

66. R. Kobayashi, *Physica D: Nonlinear Phenomena* **63**, 410 (1993).

67. N. Moelans, B. Blanpain, P. Wollants, *Calphad* **32**, 268 (2008).

68. M. Abadi, et al., arXiv preprint arXiv:1603.04467 (2016).

69. L. Q. Chen, *Annual Reviews of Materials Research* **32**, 113 (2008).

70. I. Steinbach, *Modelling and simulation in materials science and engineering* **17**, 073001 (2009).

71. S. M. Allen, J. W. Cahn, *Acta Metallurgica* **20**, 423 (1972).

72. S. M. Allen, J. W. Cahn, *Scripta Metallurgica* **7**, 1261 (1973).

73. F. Aurenhammer, R. Klein, *Handbook of computational geometry* **5**, 201 (2000).

74. A. Karma, W.-J. Rappel, *Physical review E* **57**, 4323 (1998).
Supplementary Information for
Self-supervised Learning and Prediction of Microstructure Evolution with Recurrent Neural Networks

Kaiqi Yang, Yifan Cao, Youtian Zhang, and Ming Tang
Department of Materials Science and NanoEngineering, Rice University, Houston, TX 77005, USA

Daniel Aberg, Babak Sadigh, and Fei Zhou
Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

*mt20@rice.edu
†zhou6@llnl.gov
Supplementary Figure S1: Examples of the bifurcation of grain boundary connectivity upon grain disappearance. Due to bifurcation, RNN predictions (P) differ from the ground truth (G) after the disappearance of a four-sided grain in a, or a five-sided grain in b. Top: schematics of the local topological changes that occur in predictions vs ground truth. Numbers inside each grain indicate the change to the number of sides after grain disappearance. Bottom: corresponding images from ground truth and RNN output. Red circles highlight the regions where the grain boundary connectivity bifurcates.
Supplementary Figure S2: Prediction of grain growth in a larger 256×256-pixel system using the RNN model trained on 64×64-pixel image data. a. Examples of RNN predictions (P) vs ground truth (G) from three testing cases, in which RNN outputs 200 frames based on 1 input frame. b. RMSE (black) and SSIM (blue) of the predictions averaged over 50 testing cases.
Supplementary Figure S3: RNN predictions of spinodal decomposition based on 10 input frames. a. Examples of RNN predictions (P) vs ground truth (G) from three testing cases, in which RNN outputs 200 64×64-pixel image frames based on 10 input frames, which are taken from the 41st to 50th frames of a simulation starting from a uniform mixture. b. RMSE (black) and SSIM (blue) of the predictions averaged over 510 testing cases.
Supplementary Figure S4: Evolution of the average system concentration $\langle c \rangle$ in RNN output frames. $\langle c \rangle$ is averaged over 170 testing cases with $c_0 = 0.25$ in a, 0.5 in b and 0.75 in c, respectively.
Supplementary Figure S5: Prediction of spinodal decomposition in a larger 256×256-pixel system using the RNN model trained on 64×64-pixel image data. 

a. Examples of RNN predictions (P) vs ground truth (G) from three testing cases, in which RNN outputs 200 256×256-pixel frames based on 10 input frames. 

b. RMSE (black) and SSIM (blue) of predictions averaged over 50 testing cases.
Supplementary Figure S6: Dependence of the validation error on the number of training clips $N_{\text{clip}}$ in the application of RNN to predicting plane wave propagation and grain growth.