

1      GrainNN: A neighbor-aware LSTM framework for predicting  
2      microstructure evolution during polycrystalline grain formation

3      Yigong Qin\*    Stephen DeWitt†    Balasubramanian Radhakrishnan‡    and George Biros §

4      **Supplementary Note 1: GrainNN details**

5      In Algorithm 1 of the main document, we present the GrainNN architecture with only a single NA-LSTM  
6      layer in network  $\mathcal{N}_0$  and  $\mathcal{N}$ . In Supplementary Algorithm 1, we show the full architecture with  $K$  NA-  
7      LSTM layers. The input to the first layer is the state matrix  $X_t$ ; the input to the other layers is the  
8      hidden states of the previous layer  $H_t^{k-1}$ .

9      The GrainNN hyperparameters  $J$ ,  $T$ ,  $K$ , layer size  $D_h$ , and initial learning rate are determined by  
10     grid search. The hyper-parameters producing the highest accuracy for validation runs are listed in Sup-  
11     plementary Table 1. We show training and validation loss curves for one of the models in Supplementary  
12     Fig. 1. The relative training and validation loss reduction is roughly 1E – 3 for both  $\mathcal{N}_0$  and  $\mathcal{N}$ .

13     The computational cost of GrainNN is provided in Supplementary Table 2. The inference time includes  
14     averaging predictions of the four models listed in Supplementary Table 1. For 100 in-distribution testing  
15     runs, the GrainNN-inference time is five seconds, which is 1000X faster than the GPU-accelerated phase  
16     field solver. The GrainNN-inference time costs for 100 doubled-time-horizon, doubled-initial-grain-size,  
17     and 32-grain configurations are 9, 5, and 44 s, respectively. On a NVIDIA V100 GPU, GrainNN is at  
18     least 3X faster.

**Supplementary Table 1** Hyperparameters of the top four models.

model	sequence length $J$	total snapshots $T$	# layers $K$	layer size $D_h$	initial learning rate	# parameters
1	4	20	5	32	25e-4	400034
2	4	20	4	32	50e-4	316706
3	4	20	4	32	100e-4	316706
4	6	20	4	32	100e-4	316706

19      **Supplementary Note 2: Deep-learning model comparison**

20      We compare our proposed NA-LSTM architecture with LSTM [1] and convolutional (Conv) LSTM [2].  
21      LSTM is a subclass of recurrent neural networks (RNNs) which are used for processing sequential data.

\*Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712, USA; ygqin@utexas.edu

†Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA;  
dewittsj@ornl.gov

‡Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA;  
radhakrishnb@ornl.gov

§Oden Institute, The University of Texas at Austin, Austin, TX 78712, USA; gbiros@acm.org

---

**Supplementary Algorithm 1** Architecture of GrainNN with multiple NA-LSTM layers

Input:  $X_0$ . Output:  $Z_t$  for  $t = 0, 1, \dots, T$

Hyperparameters:  $J$  is the length of input/output sequence,  $T$  is the number of time snapshots,  $K$  is the number of NA-LSTM layers.

---

```

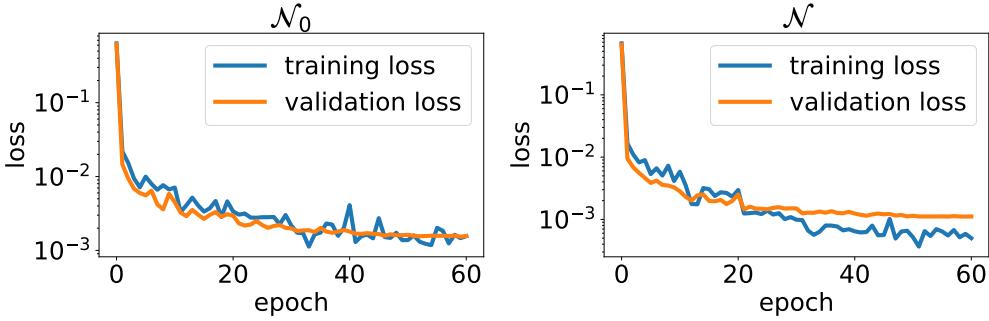
1: /*  $\mathcal{N}_0 : X_0 \rightarrow \{X_1, X_2, \dots, X_{J-1}\}$  */
2:  $\{H_0^k, C_0^k\}_{k=1}^K = 0$                                 Initialize hidden and cell states of of  $\mathcal{N}_0$ 
3: for  $t = 1 \rightarrow J - 1$  do
4:    $[H_t^1, C_t^1] = \mathcal{L}^1(X_{t-1}, [H_{t-1}^1, C_{t-1}^1])$ 
5:   for  $k = 2 \rightarrow K$  do
6:      $[H_t^k, C_t^k] = \mathcal{L}^k(H_t^{k-1}, [H_{t-1}^k, C_{t-1}^k])$ 
7:    $Z_t = \text{Output}(H_t^K)$                                 /*  $Z_t \rightarrow X_t$  using Eq. (4.4) of the main document */
8:
9: /*  $\mathcal{N} : \{X_0, X_1, \dots, X_{J-1}\} \rightarrow \{X_J, X_{J+1}, \dots, X_{2J-1}\}$  */
10:  $t^* = 0$ 
11: while  $t^* < T$  do
12:   /* Encoder:  $\mathcal{E}\{X_0, X_1, \dots, X_{J-1}\} \rightarrow \{H_{J-1}^k, C_{J-1}^k\}_{k=1}^K$  */
13:    $\{H_0^k, C_0^k\}_{k=1}^K = 0$                                 Initialize hidden and cell states of of  $\mathcal{N}$ 
14:   for  $t = 0 \rightarrow J - 1$  do
15:      $[H_t^1, C_t^1] = \mathcal{L}^1(X_{t^*+t}, [H_{t-1}^1, C_{t-1}^1])$ 
16:     for  $k = 2 \rightarrow K$  do
17:       for  $t = 0 \rightarrow J - 1$  do
18:          $[H_t^k, C_t^k] = \mathcal{L}^k(H_t^{k-1}, [H_{t-1}^k, C_{t-1}^k])$ 
19:
20:   /* Decoder:  $\mathcal{D}\{X_{J-1}, \{H_{J-1}^k, C_{J-1}^k\}_{k=1}^K\} \rightarrow \{X_J, X_{J+1}, \dots, X_{2J-1}\}$  */
21:   for  $t = J \rightarrow 2J - 1$  do
22:      $[H_t^1, C_t^1] = \mathcal{L}^1(X_{t^*+t-1}, [H_{t-1}^1, C_{t-1}^1])$ 
23:     for  $k = 2 \rightarrow K$  do
24:        $[H_t^k, C_t^k] = \mathcal{L}^k(H_t^{k-1}, [H_{t-1}^k, C_{t-1}^k])$ 
25:      $Z_t = \text{Output}(H_t^K)$                                 /*  $Z_t \rightarrow X_t$  using Eq. (4.4) of the main document */
26:    $t^* += J$ 

```

---

**Supplementary Table 2** Computational efficiency and prediction accuracy of GrainNN. PDE-solve time was measured in hours on one NVIDIA Tesla V100 GPU. GrainNN-inference time was measured in seconds on one core of Intel Xeon E5-2620 v4 CPU. Misclassification rate (MR) is an error metric of pointwise microstructure.  $L_x = 20\mu\text{m}$  and  $N_G = 8$  are the training domain size and number of grains respectively.

	training	validation	testing	2×time-horizon	2×initial-grain-size	$N_G = 32$
# runs	2400	100	100	100	100	100
PDE-solve time (h)	36	1.8	1.7	3.3	5	60
GrainNN-inference time (s)	124	5	5	9	5	44
Average MR (%)	3.7	5.7	3.7	5.4	7.6	8.3



**Supplementary Figure 1** *Training and validation losses of the second model in Supplementary Table 1.*

**Supplementary Table 3** *Performance comparison for three network architectures using the same dataset. The total number of trainable parameters is 320K for all the three network architectures.*

Network	final training loss of $\mathcal{N}$	final validation loss of $\mathcal{N}$	average misclassification rate
NA-LSTM	5.5e-4	1.1e-3	4.7%
ConvLSTM	7.0e-4	1.6e-3	6.0%
LSTM	2.6e-3	2.7e-3	12.2%

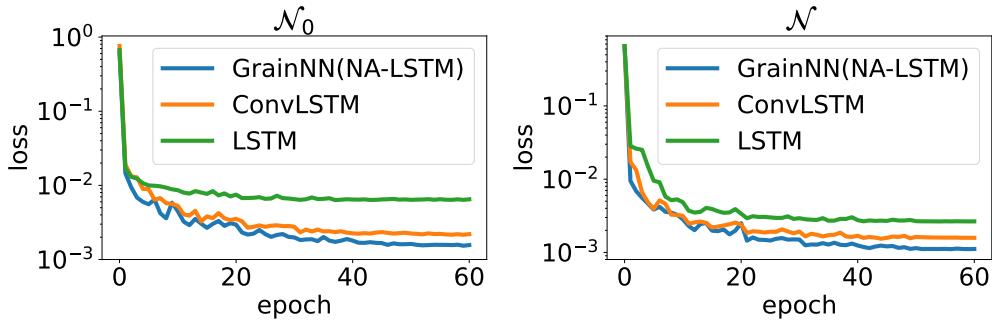
- 22 In an RNN, connections between nodes are looped, allowing information to persist within the network by  
 23 tracking an internal memory state. We use the formulation of LSTM in [3]:

$$\begin{aligned}
 i_t &= \sigma(W_i \circ [x_t, h_{t-1}] + W_{ci} c_{t-1} + b_i) \\
 f_t &= \sigma(W_f \circ [x_t, h_{t-1}] + W_{cf} c_{t-1} + b_f) \\
 c_t &= f_t c_{t-1} + i_t \tanh(W_c \circ [x_t, h_{t-1}] + b_c) \\
 o_t &= \sigma(W_o \circ [x_t, h_{t-1}] + W_{co} c_t + b_o) \\
 h_t &= o_t \tanh(C_t),
 \end{aligned} \tag{S.1}$$

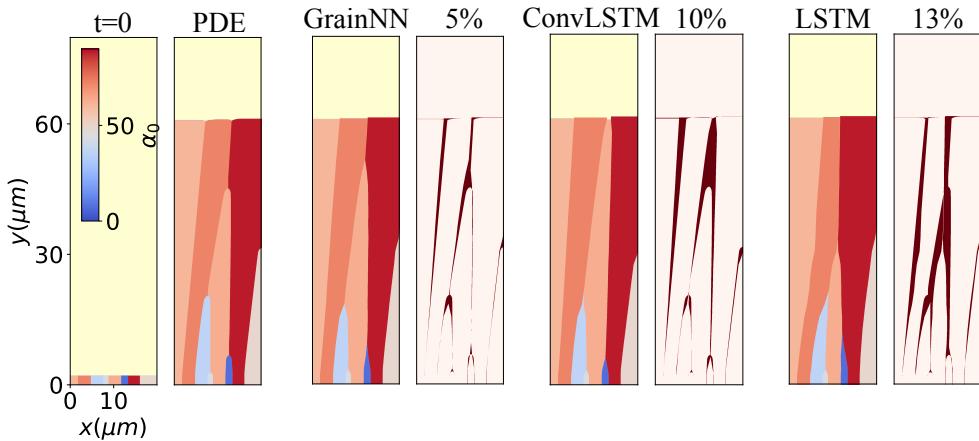
- 24 where  $i_t$ ,  $f_t$  and  $o_t$  are the input gate, forget gate, and output gate, respectively. “ $\circ$ ” represents the  
 25 matrix-vector multiplication. LSTM can retain both long-term dependencies among the input sequence  
 26 via cell states  $c_t$ , and also the short-term dependencies via hidden states  $h_t$ . For spatiotemporal predictions  
 27 where  $x_{i,t}$  is a vector defined on each location  $i$ , Eq. (S.1) is applied to every  $x_{i,t}$  independently. To  
 28 incorporate spatial correlations, ConvLSTM replaces “ $\circ$ ” with a convolution operator [2] so that the  
 29 neighboring states  $x_{i-1,t}$  and  $x_{i+1,t}$  are coupled in time predictions of  $x_{i,t}$ . In our epitaxial grain growth  
 30 setup,  $x_{i,t}$  is defined on each grain. If grain  $i+1$  is eliminated, LSTM requires a mechanism to couple  
 31  $x_{i+2,t}$  instead of  $x_{i+1,t}$ . We achieve this functionality with the devised NA-LSTM architecture, which is  
 32 an extension to the ConvLSTM. We replace “ $\circ$ ” with the attention operators so that every grain is aware  
 33 of the characteristics of its current neighbors. Supplementary Table 3 and Supplementary Fig. 2 show  
 34 that NA-LSTM outperforms ConvLSTM and LSTM with a smaller validation loss and misclassification  
 35 rate. In Supplementary Fig. 3, we provide an example of morphology prediction error for three network  
 36 architectures. GrainNN captures the grain boundary more accurately.

### 37 **Supplementary Note 3: Supplementary results for the melt pool con- 38 figuration**

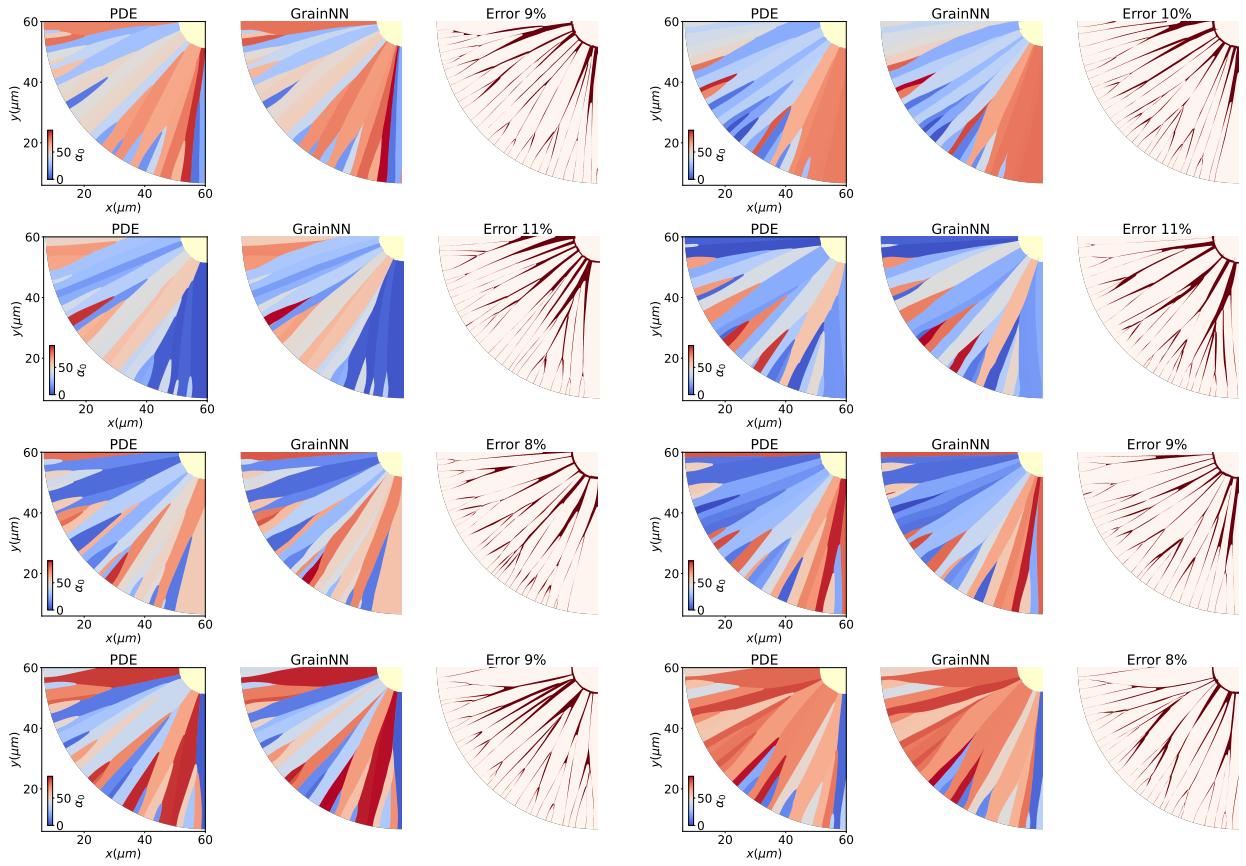
- 39 In addition to the example shown in Fig. 8 of the main document, we tested the accuracy of GrainNN on  
 40 melt-pool configuration with 100 different phase field realizations. We show eight of the realizations in



**Supplementary Figure 2 Validation loss curves for three different network architectures.** Our NA-LSTM architecture consistently outperforms the Convolutional LSTM and the vanilla LSTM. We control the total number of trainable parameters at 320K.



**Supplementary Figure 3 An example of morphology prediction error for three network architectures.** GrainNN captures grain shape more accurately compared to LSTM and ConvLSTM in presence of grain elimination, because each grain is aware of the current grain orientation of its neighbors.



**Supplementary Figure 4** *GrainNN melt-pool prediction for different initial conditions.* The physical parameters are  $G = 2.4 \text{K}/\mu\text{m}$ ,  $R = 1.52 \text{m/s}$ ,  $\epsilon_k = 0.08$ . The average misclassification rate for 100 realizations is 10.5%.

**Supplementary Table 4** Parameters comparison for Al-0.1wt%Cu, Ti-6-4, and 316L

symbol	meaning	Al-0.1 wt%Cu [4–6]	Ti-6-4 [7, 8]	316L [9, 10]
Material parameters				
$\Gamma$	Gibbs-Thompson coefficient (Km)	$2.4 \times 10^{-7}$	$4.5 \times 10^{-7}$	$3.47 \times 10^{-7}$
$k$	partition coefficient	0.14	0.838	0.791
$D_L$	diffusion coefficient in liquid ( $m^2/s$ )	$3 \times 10^{-9}$	$9.5 \times 10^{-9}$	$3 \times 10^{-9}$
$D_h$	heat diffusion coefficient ( $m^2/s$ )	$8.43 \times 10^{-5}$	$6.6 \times 10^{-6}$	$3.6 \times 10^{-6}$
$\Delta T_0$	freezing range (K)	1.6	0.17	15.7
$\mu_k$	linear kinetic coefficient ( $m/s/K$ )	0.34	0.2	0.217
$L/C_p$	latent heat/heat capacity(K)	439	258	229
$v_a$	absolute stability velocity (m/s)	0.14	0.004	0.17
phase-field-related parameters				
$W_0$	length scale (nm)	64	40	40
$\tau_0$	time scale (ns)	50	23	16
$L/(\lambda C_p)$	temperature scale (K)	4.8	11.1	14.4

41 Supplementary Fig. 4. In each subfigure, 32 grains are initialized with random grain sizes and orientations.  
42 The testing parameters are  $G = 2.4K/\mu m$ ,  $R = 1.52m/s$ ,  $\epsilon_k = 0.08$ . The average misclassification rate  
43 is 10.5%.

44 **Supplementary Note 4: Phase field solver, parameters, and results  
45 for Al-Cu.**

46 Here we provide more details of our phase field model. We use the same interface width  $W_0$  for all the  
47 order parameters. The convergence of  $W_0$  is checked with normalized grain size distributions as shown in  
48 Supplementary Fig. 5. Time scale  $\tau_0$  and thermal coupling constant  $\lambda$  are derived from matched interface  
49 asymptotic analysis [11, 12]:

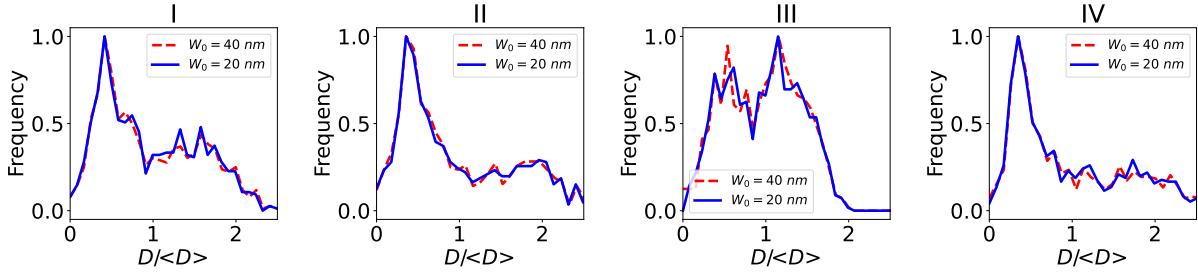
$$d_0 = a_1 \frac{W_0}{\lambda} \quad (S.2a)$$

$$\beta_0 = a_1 \frac{\tau_0}{\lambda W_0} - a_1 a_2 \frac{W_0}{D_h}, \quad (S.2b)$$

50 where  $a_1 = 5\sqrt{2}/8$ ,  $a_2 = 47/75$ ;  $d_0 = \Gamma C_p/L$  is the thermal capillarity length;  $\beta_0 = C_p/(\mu_k L)$  is the  
51 kinetic coefficient;  $D_h$  is the heat diffusion coefficient;  $\mu_k$  is the linear kinetic coefficient. For  $W_0$  we are  
52 using (Supplementary Table 4)  $W_0 \ll \beta_0 D_h$ , thus  $\tau_0 = \lambda W_0 \beta_0 / a_1$ . The obstacle parameter  $\omega$  is set to  
53  $\lambda u/I$  according to Ref. [13], where  $u$  is the nondimensional undercooling and  $I$  is a constant.

54 The discretization of the phase field equation and implementation details can be found in Ref. [14]. We  
55 use finite difference discretization in space and forward Euler in time. For simulations in the rectangular  
56 domain, we apply the moving-domain technique [15] to save cost. The performance and scaling of  
57 the phase field solver is available at [https://github.com/YigongQin/PF\\_cpp](https://github.com/YigongQin/PF_cpp). The code is highly  
58 optimized with GPU. We use CUDA-aware MPI which scales well on multiple GPUs and multiple nodes.

59 The adopted phase field model [7] is valid for rapid solidifications, e.g., additive manufacturing,  
60 where the interface is planar and there are no cellular and dendritic structures present [16–18]. In this  
61 regime, the growth rate of directional solidification is higher than the absolute stability velocity  $V_a$  which



**Supplementary Figure 5** *Convergence check of phase-field simulations for the four sets of testing physical parameters.* The blue lines are normalized grain size distributions for interface width  $W_0 = 20\text{ nm}$ ; the red lines are normalized grain size distributions for interface width  $W_0 = 40\text{ nm}$ . We use 100 realizations for each set of physical parameter. The results show that  $W_0 = 40\text{ nm}$  is small enough for phase field simulations.

**Supplementary Table 5** Videos for out-of-distribution generalizations. The training configuration is eight-grain and rectangular geometry.

video	physical parameters	description
target_rectangular.mp4	$G = 2.4K/\mu\text{m}$ , $R = 1.52\text{m/s}$ , $\epsilon_k = 0.08$	Fig. 6c of the main document, 32 grains
target_meltpool.mp4	$G = 2.4K/\mu\text{m}$ , $R = 1.52\text{m/s}$ , $\epsilon_k = 0.08$	Fig. 8 of the main document, 32 grains
128grain.mp4	$G = 4.75K/\mu\text{m}$ , $R = 1.82\text{m/s}$ , $\epsilon_k = 0.27$	128 grains

62 is obtained by [19]:

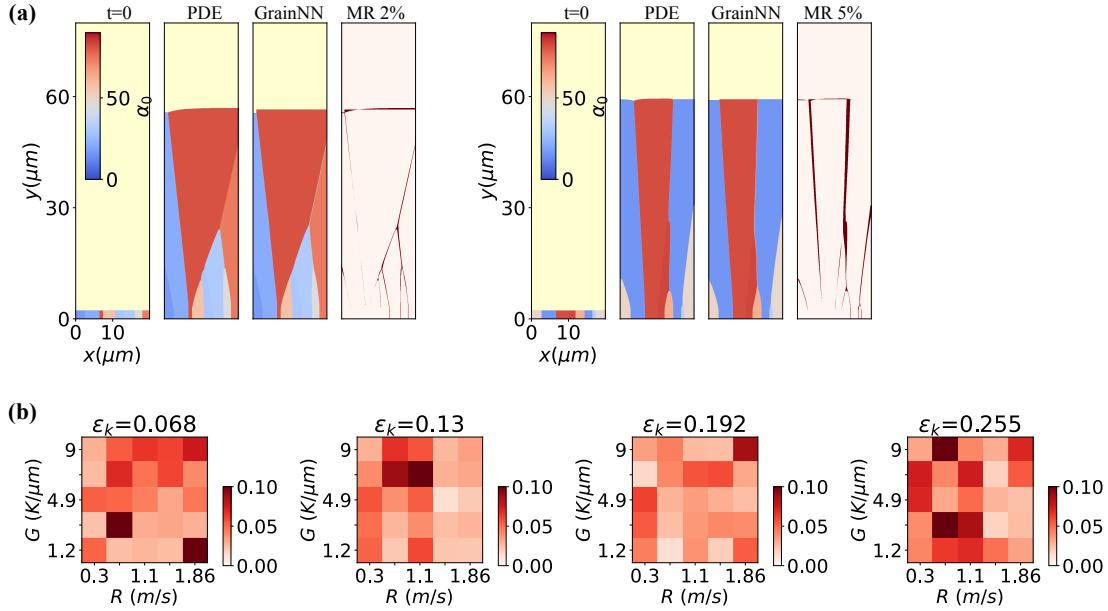
$$V_a = D_L \Delta T_0 / \Gamma k, \quad (\text{S.3})$$

63 where  $D_L$  is the diffusion coefficient in liquid;  $\Delta T_0$  is the equilibrium freezing range;  $k$  is the distribution  
64 coefficient;  $\Gamma$  is the Gibbs-Thompson coefficient.  $V_a$  for three materials stainless steel 316L, Ti-6Al-4V,  
65 and Al-0.1 wt%Cu are shown in Supplementary Table 4. For 316L,  $V_a = 0.17\text{m/s}$ , which is comparable  
66 to 0.12m/s predicted by Pinomaa et al. [10] and 0.093m/s predicted by Chadwick et al. [9]. We can see  
67 that for pulling velocity 0.2–2m/s investigated in this paper, the phase field model is valid for all three  
68 alloys.

69 In the main document we show GrainNN results for 316L. We did complementary experiments to show  
70 GrainNN can be used to predict microstructures of Al-Cu as well. We collected the same amount of phase  
71 field data with Al-Cu and retrained GrainNN. We test GrainNN with the same physical parameters used  
72 for 316L. The pointwise error is examined with the misclassification rate (MR) shown in Supplementary  
73 Fig. 6. The average MR is 4.5% compared to 3.7% for 316L.

## 74 Supplementary Note 5: Videos

75 We provide videos (Supplementary Table 5) of the microstructure evolution for three out-of-distribution  
76 generalization configurations. In each video, we show the high-fidelity phase field simulation, the GrainNN  
77 prediction, and the pointwise error. The first video corresponds to the 32-grain configuration in Fig. 6c  
78 of the main document. The second video corresponds to the melt-pool configuration in Fig. 8 of the  
79 main document. We show another video with 128 grains. The final misclassification rate of the 128-grain  
80 case is 10%.



**Supplementary Figure 6** *GrainNN accuracy for Al-Cu.* (a) Two examples from the testing simulations. (b) Misclassification rates (MR) at the final time. The 100 sets of testing parameters include five different  $G$ , five different  $R$ , and four different  $\varepsilon_k$ . Each point in the error heatmap corresponds to MR of one phase-field simulation. The average MR is 4.5%.

## 81 Supplementary References

- 82 [1] Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural computation* **9**, 1735–1780  
83 (1997).
- 84 [2] Xingjian, S. et al. Convolutional lstm network: A machine learning approach for precipitation  
85 nowcasting. In *Advances in neural information processing systems*, 802–810 (2015).
- 86 [3] Graves, A. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*  
87 (2013).
- 88 [4] Nath, S. K. D., Shibuta, Y., Ohno, M., Takaki, T. & Mohri, T. A molecular dynamics study of  
89 partitionless solidification and melting of Al–Cu alloys. *Isij International* **57**, 1774–1779 (2017).
- 90 [5] Farzadi, A., Do-Quang, M., Serajzadeh, S., Kokabi, A. H. & Amberg, G. Phase-field simulation of  
91 weld solidification microstructure in an Al–Cu alloy. *Modelling and Simulation in Materials Science  
92 and Engineering* **16** (2008).
- 93 [6] Takaki, T., Ohno, M., Shimokawabe, T. & Aoki, T. Two-dimensional phase-field simulations of  
94 dendrite competitive growth during the directional solidification of a binary alloy bicrystal. *Acta  
95 Materialia* **81**, 272–283 (2014).
- 96 [7] Pinomaa, T. et al. Process-Structure-Properties-Performance modeling for selective laser melting.  
97 *Metals* **9**, 1138 (2019).
- 98 [8] Sahoo, S. & Chou, K. Phase-field simulation of microstructure evolution of Ti–6Al–4V in electron  
99 beam additive manufacturing process. *Additive manufacturing* **9**, 14–24 (2016).

- 100 [9] Chadwick, A. F. & Voorhees, P. W. The development of grain structure during additive manufac-  
101 turing. *Acta Materialia* **211**, 116862 (2021).
- 102 [10] Pinomaa, T., Lindroos, M., Walbrühl, M., Provatas, N. & Laukkanen, A. The significance of spatial  
103 length scales and solute segregation in strengthening rapid solidification microstructures of 316L  
104 stainless steel. *Acta Materialia* **184**, 1–16 (2020).
- 105 [11] Karma, A. & Rappel, W.-J. Quantitative phase-field modeling of dendritic growth in two and three  
106 dimensions. *Physical review E* **57**, 4323 (1998).
- 107 [12] Bragard, J., Karma, A., Lee, Y. H. & Plapp, M. Linking phase-field and atomistic simulations to  
108 model dendritic solidification in highly undercooled melts. *Interface Science* **10**, 121–136 (2002).
- 109 [13] Ofori-Opoku, N. & Provatas, N. A quantitative multi-phase field model of polycrystalline alloy  
110 solidification. *Acta Materialia* **58**, 2155–2164 (2010).
- 111 [14] Qin, Y., Bao, Y., DeWitt, S., Radhakrishnan, B. & Biros, G. Dendrite-resolved, full-melt-pool phase-  
112 field simulations to reveal non-steady-state effects and to test an approximate model. *Computational  
113 Materials Science* **207**, 111262 (2022).
- 114 [15] Badillo, A. & Beckermann, C. Phase-field simulation of the columnar-to-equiaxed transition in alloy  
115 solidification. *Acta Materialia* **54**, 2015–2026 (2006).
- 116 [16] Mullins, W. W. & Sekerka, R. Stability of a planar interface during solidification of a dilute binary  
117 alloy. *Journal of applied physics* **35**, 444–451 (1964).
- 118 [17] Merchant, G. & Davis, S. Morphological instability in rapid directional solidification. *Acta metal-  
119 lurgica et materialia* **38**, 2683–2693 (1990).
- 120 [18] Trivedi, R. & Kurz, W. Morphological stability of a planar interface under rapid solidification  
121 conditions. *Acta metallurgica* **34**, 1663–1670 (1986).
- 122 [19] Trivedi, R., Sekhar, J. & Seetharaman, V. Solidification microstructures near the limit of absolute  
123 stability. *Metallurgical Transactions A* **20**, 769–777 (1989).