Optimization Analysis of a Deep Learning-Based Model for Predicting Temperature Fields in the Solidification Process of Castings

Yahui Yang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University

Beijing, Haidian District, China

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University

Beijing, Haidian District, China

Jinwu Kang

School of Materials Science and Engineering, Tsinghua University

Beijing, Haidian District, China
E-mail: 22126065@bjtu.edu.cn,jwwang@bjtu.edu.cn,kangjw@tsinghua.edu.cn

Abstract

In the foundry industry, accurate prediction of the temperature field of castings during solidification is crucial to ensure high product quality and improve productivity. This study provides an in-depth investigation of the Unet deep learning model used for temperature field prediction during solidification, focusing on the impact of key training parameters such as optimiser selection, batch size, number of iterations, and loss function selection on the prediction performance of the Unet model. The results of the study demonstrate the analysis of selecting and tuning these parameters to improve the accuracy and reliability of model predictions. The results of this study not only provide new ideas for the practice of deep learning applications in the foundry industry, but also help to improve the accuracy and efficiency of the production proces.

Keywords: casting, deep learning, temperature field prediction, parameter optimization.

1. Introduction

In recent years, deep learning has expanded in field prediction, simplifying traditional numerical simulations [1], [2] while maintaining accuracy. Specifically, in casting solidification, data-based methods predict temperature fields [3] effectively. However, incorrect network structures or training parameters can cause underfitting or overfitting, leading to errors. Thus, optimizing these parameters is vital for enhancing accuracy, reducing costs, and speeding up design iterations for precise temperature field prediction in casting processes.

2. Model Introduction

Selecting an appropriate network model structure is crucial for high-precision prediction of the temperature field during the solidification process of castings. The U-Net [4], [5] network demonstrates exceptional performance in the field of image segmentation, particularly in achieving precise segmentation with smaller datasets.

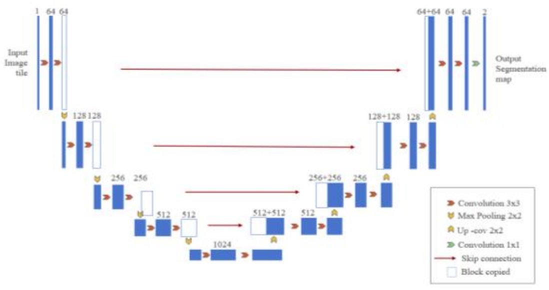

Fig.1. Unet network model [6]

In the prediction of temperature field during casting solidification, due to the time-consuming computation of the finite difference method, it is difficult to generate a large amount of training data, and the encoder-decoder network architecture is able to efficiently extract the semantic features of the image and retain the positional information, which is suitable for prediction tasks with small datasets. The Unet network model is shown in Fig. 1.

In this study, we employ a network model (Fig. 1) with the classic U-Net structure to predict the temperature field during the solidification process of castings. The model takes the mold images of sand castings as input, and the output is the temperature field image after a certain time step. Both input and output are single-channel images.

3. Optimization Methods

In the realm of deep learning network architectures, the addition or alteration of various modules can significantly impact the final predictive performance. To achieve optimal neural network prediction capabilities, this study delves into several key factors affecting model training, including the choice of optimizer, batch size, number of iterations, and the selection of loss functions. The following sections will present a detailed comparative analysis of these crucial elements, aiming to elucidate how to optimize deep learning network structures for higher accuracy in predictions. This translation is intended for application in academic papers.

3.1. *The Impact of Different Optimizers on the Prediction Accuracy of the Temperature Field*

In this set of comparative experiments, the batch size was set to 32, and the number of epochs was fixed at 200. Mean Absolute Error(MAE) was chosen as the loss function. To analyze the impact of different optimizers, three distinct optimizers were selected for comparison: Adam, Nadam and SGD. The learning rate for each was adjusted using a cosine annealing learning rate strategy. The experimental results are presented in Table 1.

Table 1. Comparison of results of different optimizers

Serial Number	Optimizer	Accuracy	Pearson Correlation Coefficient	Loss	Val-Loss	Training Time
1	Adam	0.9156	0.9989	0.0096	0.0230	66727.5657
2	SGD	0.8198	0.7685	0.1977	0.1769	68961.8629
3	Nadam	0.9118	0.9989	0.0099	0.0236	65335.4521

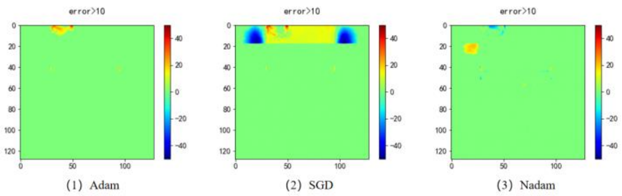

Fig.2. Error Plots of Predicted versus Actual Values Using Three Optimizers

According to the data presented in Table 1, the model demonstrates higher accuracy and correlation when employing the Adam optimizer, with the smallest loss value observed. To further analyze and compare these results, samples at a specific moment (t=10 minutes) were randomly selected for visual analysis. Fig.2 demonstrates the distribution of errors between the temperature fields predicted by the model and the actual temperature fields, under three different optimizers: Adam, Nadam, and SGD.

3.2. *The Impact of Different Batch Sizes on Prediction Accuracy*

In this set of comparative experiments, building upon the results of the previous set, Adam was selected as the optimizer, with the number of epochs set to 200 and MAE employed as the loss function. To investigate the impact of batch size on model performance, three different batch sizes (16, 32, and 64) were established for comparative analysis. The experimental results are shown in Table 2.

Table 2. Comparison of results for different batch sizes (batch size)

Serial Number	Optimizer	Accuracy	Pearson Correlation Coefficient	Loss	Val-Loss	Training Time
1	16	0.9014	0.9889	0.0181	0.0217	68537.3182
2	32	0.9118	0.9989	0.0096	0.0230	66727.5657
3	64	0.8089	0.9979	0.0214	0.0286	66357.1713

Based on the data analysis from Table 2, it is observed that setting the batch size to 32 results in higher model accuracy and correlation, with the smallest loss value and shorter training time. Samples at the same moment (t=10 minutes) were selected for visual analysis.

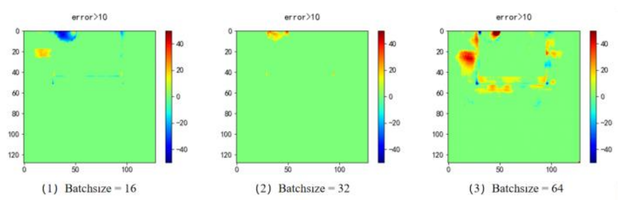

Fig.3. Error Plots of Predicted versus Actual Values for Three Batch Sizes

Fig. 3 illustrates the distribution of errors between the predictedand actual values of the model under three different batch sizes (16, 32, and 64). The model exhibits relatively smaller prediction errors when the batch size is set to 32.

3.3. *The Impact of Different Iteration Counts on Prediction Accuracy*

In this set of comparative experiments, building upon the results of the first two sets, Adam was chosen as the optimizer, with the batch size set to 32 and Mean Absolute Error (MAE) used as the loss function. To explore the impact of different numbers of epochs on model performance, three distinct numbers of epochs (50, 200, and 500) were established for comparative analysis. The experimental results are shown in Table 3.

According to the experimental data in Table 3, the model exhibits the highest accuracy and correlation, with the smallest loss value and relatively short training time, when the number of epochs is set to 200. With an increase in the number of epochs, the computational cost correspondingly rises. Therefore, when selecting the optimal parameters for the model, it is crucial to consider both prediction accuracy and computational cost as key factors in determining the most suitable parameter settings. Fig. 4 displays error distribution of the temperature field at different numbers of epochs (50, 200, 500), providing a more intuitive assessment of model performance.

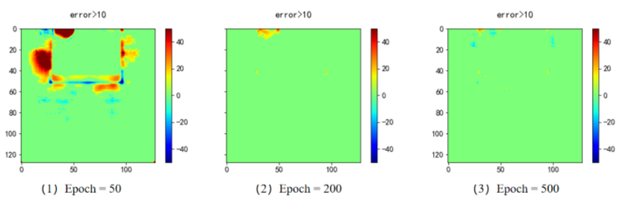

Fig.4. Error Plots of Predicted versus Actual Values for Three Different Epoch Counts

Table 3. Comparison of results for different number of iterations (epoch)

Serial Number	Epoch	Accuracy	Pearson Correlation Coefficient	Loss	Val-Loss	Training Time
1	50	0.7264	0.8963	0.0302	0.0361	17488.4049
2	200	0.9118	0.9989	0.0096	0.0230	66727.5657
3	500	0.9083	0.9988	5.3330e-04	0.0040	157621.8659

3.4. *The Impact of Different Loss Functions on Prediction Accuracy*

In this set of comparative experiments, based on the results of the previous three sets, we selected the Adam optimizer with the number of epochs set to 200 (epoch = 200). Three different loss functions were chosen for comparative analysis: Mean Squared Error (MSE), Mean Absolute Error (MAE), and their combination (MSE + MAE). The experimental results are presented in Table 4.

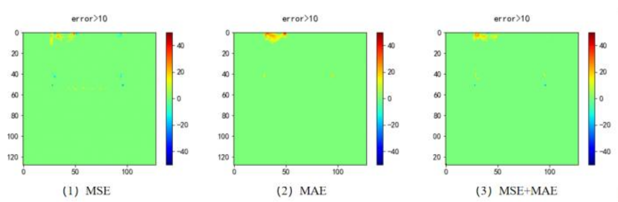

Fig.5. Error Plots of Predicted versus Actual Values for Three Different Loss Functions

Based on the experimental data presented in Table 4, the highest accuracy and correlation, along with the smallest

Table 4. Comparison of results of different loss functions

Serial Number	Loss Function	Accuracy	Pearson Correlation Coefficient	Loss	Val-Loss	Training Time
1	MSE	0.9293	0.9990	4.9873e-04	0.0038	64514.8804
2	MAE	0.9118	0.9989	0.0096	0.0230	66727.5657
3	MAE+MSE	0.9083	0.9988	5.3330e-04	0.0040	67621.8659

loss value and relativly short training time, are observed when Mean Squared Error (MSE) is used as the loss function. Thus, choosing MSE as the loss function is the optimal parameter choice for the model. Fig. 5 illustrates the visualization of the error distribution of the temperature field predictions when different loss functions (MSE, MAE, and MSE+MAE) are employed.

3.5. *Experimental Results*

In the four comparative experiments conducted, we determined that the Adam optimizer is the optimal choice, surpassing Nadam and SGD in accuracy, correlation, and loss minimization. The model exhibits its best performance with a batch size of 32, effectively balancing efficiency and accuracy. Setting the number of epochs to 200 achieves a balance between computational cost and precision. Among the loss functions compared, Mean Squared Error (MSE) significantly excelled in predictive accuracy and minimizing loss.

4. Conclusion

In this study, we employed a data-driven Unet neural network model for temperature field prediction, optimizing its structure and training parameters (optimizer, batch size, epochs, loss function). The model, trained on 128x128 pixel image data, showed optimal performance with the Adam optimizer, a batch size of 32, 200 epochs, and Mean Squared Error (MSE) as the loss function, achieving an accuracy of 0.9293 and a correlation of 0.9990. These findings offer new insights and valuable references for the application of deep learning in the casting industry.

Reference

1. Lee S M, Lee W J. Finite-element analysis on thermornechanical behavior of a marine propeller casting in the sand-casting process[J]. JOURNAL OF MATERIALS ENGINEERING AND PERFORMANCE, 2005, 14(3): 388-394.
2. Liu S, Bai L, Wang B, el. Numerical simulations of solidification and hot tearing for continuous casting of duplex stainless steel[J]. JOURNAL OF IRON AND STEEL RESEARCH INTERNATIONAL, 2020, 27(6): 643-655.
3. Cai S, Wang Z, Wang S, et al. Physics-Informed Neural Networks for Heat Transfer Problems[J]. Journal of Heat Transfer, 2021, 143(6): 060801.
4. Gao W, Zhao J, Peng S. UNet–Based Temperature Simulation of Hot Dry Rock in the Gonghe Basin[J]. Energies, 2022, 15(17): 6162.
5. Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686-707.
6. Han xiao. Research on Temperature Field Prediction Method of Sand Mold Casting Solidification Process Based on U-NET [D]. Beijing Jiaotong University,2022

Authors Introduction

Ms. Yahui Yang

She is currently pursuing her master's degree in Mechanical Engineering at Beijing Jiaotong University. She is mainly engaged in the research on deep learning analogue simulation.

Dr. Jiwu Wang

He is an associate professor, at Beijing Jiaotong University. His research interests are Intelligent Robots, Machine Vision, and Image Processing.

Dr. Jinwu Kang

He is an associate researcher and doctoral supervisor at Tsinghua University, specializing in the fields of intelligent casting and simulation modeling.