

Predicting Heat Transfer Coefficient Using Bidirectional Long Short-Term Memory

Ankan Basu¹, Aritra Saha², Sumanta Banerjee³

¹ Department of Computer Science and Engineering, Jadavpur University, Kolkata, India

²Department of Computer Science and Engineering, Heritage Institute of Technology, Kolkata, India

³Department of Mechanical Engineering, Heritage Institute of Technology, Kolkata, India

ABSTRACT

One of the most important steps in improving the performance of thermal systems during heat treatment processes is the estimation of the Heat Transfer Coefficient (HTC). Traditional numerical methods encounter challenges when solving the Inverse Heat Transfer Problem (IHTP) for HTC prediction. This study introduces an innovative machine learning approach, utilizing the capabilities of neural networks - specifically Bidirectional Long Short-Term Memory (BiLSTM) networks, to predict HTC values. The constructed model achieves remarkable precision with an accuracy of $\approx 98.75\%$, surpassing conventional feed-forward networks. This approach offers multiple advantages, including swift estimations of the essential features of the HTC function, delivering prompt insights into the heat transfer process. The findings underscore the potential of machine learning in optimizing heat treatment processes and sets the stage for further research that apply more sophisticated machine learning algorithms and look into other factors that affect the HTC.

Keywords: Heat Transfer Coefficient Prediction, Inverse Heat Transfer Problem, Machine Learning, Bidirectional Long Short-Term Memory, Neural Networks.

NOMENCLATURE

ANN	Artificial Neural Network
BiLSTM	Bidirectional Long Short Term Memory
CNN	Convolutional Neural Network
FFNN	Feed Forward Neural Network
HTC	Heat Transfer Coefficient ($\text{W/m}^2\text{K}$)
IHTP	Inverse Heat Transfer Problem
LSTM	Long Short Term Memory
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MSE	Mean Squared Error
PINN	Physics Informed Neural Network

1. INTRODUCTION

Heat Transfer Coefficient (HTC) is a measure of the rate of transfer of thermal energy per unit temperature difference across unit area. It is an indication of how efficiently thermal energy is transferred across the boundary of two surfaces; e.g. the boundary between the metal surface and water during quenching of a hot metal piece. Thus, it becomes very important to determine the values of the HTC, during heat treatment, in order to optimise thermal performance and tune the mechanical properties of a material to suit specific applications.

Literature survey indicates that the microstructure of materials can be customized for various industrial applications by the use of heat

treatment techniques (Oksman et. al. [1]; Miłkowska-Piszczyk and Falkus [2]). In general, metallurgical properties of cast metals can be affected by their cooling rates during solidification or on thermal treatment. Secondary cooling conditions are also critical to both process efficiency and product quality in continuous casting applications. The rate of cooling is an important control parameter; as thermal stresses can be potentially generated inside the product by extremely rapid cooling rates, leading to fractures or cracks. On the other hand, when a specific casting process is exposed to extremely low cooling rates, its metallurgical and economic advantages may be compromised. Due to these reasons, developing a basic understanding of the trends in the variance of (local and/or average) HTC values related to different cooling methods becomes crucial in industrial practice.

The computation of cooling rates can be facilitated by this understanding of the trends in variation of HTC values, thereby enabling the assessment of the suitability of these different cooling techniques for a manufacturing process. The main goal of the present work is to use machine learning techniques to estimate the (local) HTC values from hot metal surfaces to the surrounding coolant (water). Thus, the practical significance of this paper is justified in industrial settings where continuous metal casting is performed (Bamberger and Prinz [3]; Ramírez-López et al.[4]). From the above discussions, it is evident that estimating the (average) HTC values is

the primary challenge in accomplishing desirable micro-structural changes in a workpiece.

The HTC prediction is an Inverse Heat Transfer Problem (IHTP). It cannot be easily solved using traditional numerical methods due to various factors such as non-uniqueness of the solution, ill-posedness, complex nature of physical systems, nonlinearities, limited observations, etc.

Many strategies have been used to address this problem. These include heuristic search algorithms like Particle Swarm Optimisation (Carlos et al. [5]; Vakili and Gadala [6]), Genetic Algorithms (Colaco et al. [7]; Raudenský et al. [8]; Özisik and Orlando [9]), and other meta-heuristic algorithms (Sun et al. [10]; Carlos et al. [5]). However, these techniques often have large processing requirements (Carlos et al. [5]).

Alternatively, prediction of HTC can be done by using machine learning techniques, offering rapid approximations of the essential elements of the HTC function. Machine learning approaches are appropriate for HTC estimating tasks, as complicated correlations in the data can be efficiently captured by them (Anderson et al. [11]).

In recent years machine learning techniques have been employed by several researchers to solve the IHTP. They are briefly discussed below.

The ANN has been used by several researchers including Sreekanth et al. [12], Soeiro et al. [13], Mirsephai et al. [14], Patel et al. [15] to solve IHTP.

The solution of forward and inverse steady-state heat conduction problems in multilayer media, using an innovative type of PINN, has been reported by Zhang et al. [16].

The LSTM-based encoder-decoder neural network to estimate online heat flux at the tool-chip region, during turning process, has been employed by Han et al. [17].

The real-time prediction of multi-dimensional thermal boundary condition parameters using a combined model of CNN and LSTM has been proposed by Zhu et al. [18].

The estimation of HTC values in immersion quenching process was done using FFNN by Szénási and Felde [19], and promising results were reported. However dissatisfaction over the accuracy of their method and the need for improvement was expressed by the authors. Unfortunately, this is quite time-consuming due to large amounts of data. Thus, as a remedial measure, the database has been made available by the authors to the global research community, allowing for collaboration and the investigation of potential fixes.

In the present work, an attempt is made to extend the work of Szénási & Felde [19] and to find better neural network architecture for the dataset. Investigations reveal that the BiLSTM-based model outperform the other models and

produce excellent accuracy of predicting the variations of (local) values of the HTC.

2. METHODOLOGY

2.1 Data

For the present study, the HTC dataset developed and reported in [19] has been used.

Data of water quenching process of a cylindrical Inconel 600 bar (of 60 mm length and 20 mm diameter) from 850°C to room temperature is presented in the dataset. Temperature data was recorded for 1 minute at an interval of 0.5 sec. So, each temperature data point is represented as a vector of 120 elements. After the necessary preprocessing, each HTC data point has been represented as a vector of 120 elements corresponding to each temperature value recorded at the 0.5 sec intervals.

A sample data from the data set is shown in Fig. 1. In the figure, the HTC has been plotted on the Y axis and temperature on the X axis.

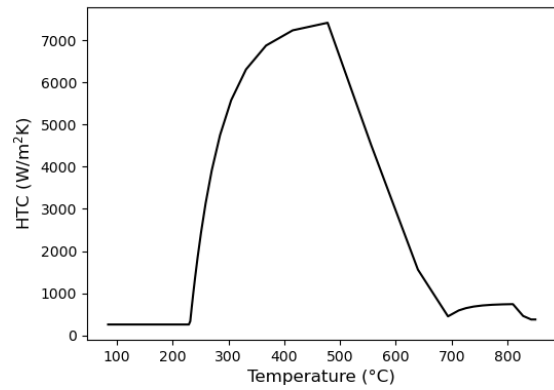


Fig. 1: A sample data from the dataset (HTC versus Temperature)

The training data set consists of 1,000,000 data points while the validation and testing data sets consists of 100,000 data points each.

2.2 Model

In this study, experimentation was done with different model architectures. The best fit model is based on BiLSTM (refer to Fig. 2). It has three layers of BiLSTM containing 2×50, 2×100 and 2×100 nodes (in the order from input layer to output layer). At the end there is a normal feed forward layer with 200 input and 1 output node. The last layer has a linear activation function.

The loss function was chosen as the MSE, and the Adam optimiser was used to train the network. The MAPE was used to evaluate the performance of the model. Training was done in batches of size 32, as it is shown to be in the range of optimal batch size for deep neural network training by Masters and Luschi [20].

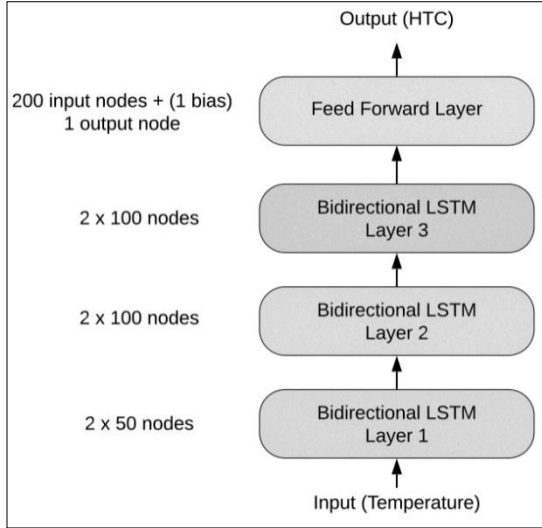


Fig. 2: Schematic representation of the proposed model

2.3 Training Environment

The training the models have been done with TensorFlow 2.12.0 in Python 3.10.10 on T4 GPU.

3. RESULTS AND DISCUSSIONS

High accuracy ($\approx 1.25\%$ error) has been yielded by the BiLSTM model. Very similar values of the MAPE are exhibited by the training, validation and testing datasets, indicating good generalisability and absence of overfitting.

To assess the efficacies of different ML methodologies, similar models based on RNN and LSTM (with same architecture as the proposed best model) were trained and evaluated. In addition, a simple FFNN with two hidden layers (50 nodes, 100 nodes) and one output layer of 120 nodes was also trained and evaluated.

The MAPE values is shown in Table 1.

Table 1: MAPE values for our proposed model

Model	Training	Validation	Testing
FFNN	8.6321	8.6383	8.6214
RNN	5.8751	5.8810	5.8768
LSTM	2.3431	2.3457	2.3425
BiLSTM	1.2529	1.2572	1.2548

The performance of recurrent models (RNN, LSTM, and BiLSTM) is found to significantly surpass that of the simple FNN model. This superiority can be attributed to the recurrent models' capability to capture temporal relationships within the data. The notably higher performance exhibited by the LSTM-based models, as compared to the simple RNN model, is due to their effective utilization of long and short-term memory mechanisms, enabling them to retain relevant information in long temporal data. The exceptional

results demonstrated by the BiLSTM model can be attributed to their proficiency in extracting context from both past and future information in temporal data, which has made them the best model in this numerical investigation to model the complex relationship between the values of HTC with temporal temperature data.

A plot of the HTC values versus temperature for a data point in the testing data set is shown in Fig. 3. Both the original and predicted values (by the BiLSTM model) are shown, and it can be seen that they are very close, indicating good prediction.

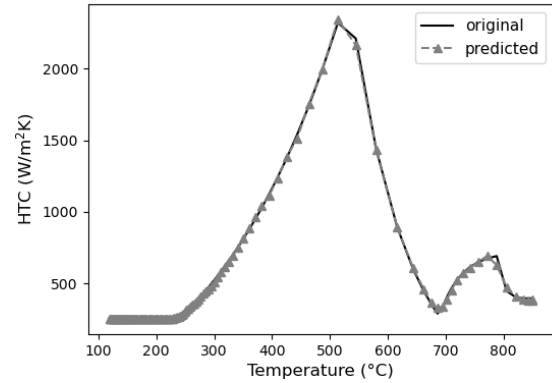


Fig. 3: Comparison of predicted versus original HTC value for a data point in the test dataset

4. CONCLUSIONS

This study presents a ML based approach to solve the inverse heat transfer problem to overcome the hurdles faced in using traditional numerical methods to solve ill-posed problems. It outlines the capability of the ML algorithms, specifically the BiLSTM to map the complicated relationship between temperature (field variable) and HTC (thermal parameter) in case of one dimensional IHTP. Future research scopes may include investigating more complex heat transfer scenarios and using more advanced neural network architectures such as encoder-decoder, PINNs, etc.

5. REFERENCES

- [1] Oksman, P., Yu, S., Kytönen, H., and Louhenkilpi, S., 2014, Acta Polytech. Hung., 11, 5-22.
- [2] Miłkowska-Piszczyk, K., and Falkus, J., 2014, Metalurgija, 53(4), 571-573.
- [3] Bamberger, M., and Prinz, B., 1986, Mater. Sci. Technol., 2(4), 410-415.
- [4] Ramírez-López, A. A.-L., Palomar-Pardavé, M., Romero-Romo, M. A., and Muñoz-

- Negrón, D., 2010, Int. J. Miner. Metall. Mater., 17, 403-416.
- [5] Carlos, A., Coello, C., David, A., Van, V., and Gary B., L., 2013, *Evolutionary Algorithms for Solving Multi-Objective Problems*, Second Edition, Springer.
- [6] Vakili, S., & Gadala, M. S., 2009, Numer Heat Tr B-Fund, 56(2), 119-141.
- [7] Colaco, M., Orlande, H., and Dulikravich, G., 2006, J Braz Soc Mech Sci., 28, 1-24.
- [8] Raudenský, M., Woodbury, K. A., Kral, J., and Brezina, T., 1995, Numer Heat Tr B-Fund, 28(3), 293-306.
- [9] Özisik, M. N., and Orlande, H. R., 2000, *Inverse Heat Transfer: Fundamentals and Applications*, Second Edition, CRC Press.
- [10] Sun, S.-C., Qi, H., Ren, Y.-T., Yu, X.-Y., and Ruan, L.-M., 2017, Int. Commun. Heat Mass Transf., 87, 132-146.
- [11] Anderson, C. W., Hittle, D. C., Katz, A. D., and Kretchmar, R. M., 1997, Artif. Intell. Eng., 11(4), 421-429.
- [12] Sreekanth, S., Ramaswamy, H., Sablani, S., and Prasher, S., 1999, J. Food Process. Preserv., 23(4), 329-348.
- [13] Soeiro, F., Soares, P. O., Campos Velho, H., and Silva Neto, A., 2004 Inverse Problems, Design and Optimization Symposium - IPDO-2004. Rio de Janeiro, Brazil, 2004, 358-363.
- [14] Mirsephai, A., Mohammadzaheri, M., Chen, L., and O'Neill, B., 2012, Int. Commun. Heat Mass Transf., 39(1), 40-45.
- [15] Patel G. C. M, Shettigar, A. K, Krishna, P. and Parappagoudar, M. B, 2017, Appl. Soft Comput., 59, 418-437.
- [16] Zhang, B., Wu, G., Gu, Y., Wang, X., and Wang, F., 2022, Phys. Fluids, 34, 116116.
- [17] Han, J., Xu, L., Cao, K., Li, T., Tan, X., Tang, Z. and Liao, G., 2021, Case Stud. Therm. Eng., 26, 101002.
- [18] Zhu, F., Chen, J., Han, Y., & Ren, D., 2022, Int. J. Heat Mass Transf., 194, 123089.
- [19] Szénási, S., Felde, I., 2019, Data, 4, 90.
- [20] Masters, D., and Luschi, C., 2018, arXiv preprint arXiv:1804.07612.