CNN-LSTM-based Study on the Dynamic Characteristics of Steel Protection Slag Crystallization

Sisi Dong[1], Zilong Cheng[*], Ziwei Cheng[2]

{352348420@qq.com[1], 2253018725@qq.com[*], 2254442311@qq.com[2]}

School of Computer and Information Engineering, Hubei Normal University, Huangshi 435000, China

Abstract: To improve steel production efficiency and product quality, enhance market competitiveness, secure the development of the steel industry, and reduce the high cost and loss brought by manual operations to the steel industry, this paper proposes a CNN-LSTM-based model for studying the dynamic characteristics of protection slag crystallization in steel. For the accuracy and practicality of protection slag crystallization prediction, the dynamic characteristics of protection slag crystallization are studied at the data level, the CNN-LSTM network model is built and the LSTM model is used to capture the temporal information of the behaviour and to predict the protection slag crystallization process. The experimental results show that the CNN-LSTM-based model can better analyse and predict the crystallization behaviour using sequence images. Compared with the traditional supervised and unsupervised learning methods, the model has higher accuracy and practicality.

Keywords: steel, protective slag, crystallization behavior, dynamic features, CNN-LSTM, feature sequences

1 INTRODUCTION

The melting and crystallization of protective slag is a key step in the steel smelting process[1]. Crystallization is the process of forming crystals through the gradual solidification of dissolved substances, and at high temperatures over a long period of time, certain crystalline substances may lose their stability and the crystal structure may change, leading to the rearrangement of atoms or molecules within them and the formation of defects or impurities. So it is necessary to cool down the temperature in time after crystallization. Analysis of the crystallization function of protection slag is of great significance for determining the quality of steel products. However, at present, the analysis of the crystallization behaviour of protection slag still requires a lot of manual operations, which not only wastes manpower and increases production costs, but also easily leads to over-crystallization and waste. And steel as one of the important pillars of modern economy, improving the production efficiency and product quality has become an urgent problem to be solved.

For behavioural research, the three main types of common approaches include supervised learning, unsupervised learning, and deep learning[2]. Compared with supervised and unsupervised learning, deep learning can handle more complex tasks, can provide a more comprehensive understanding of behaviour, and can use useful labelling information to improve

ICIDC 2023, June 02-04, Nanchang, People's Republic of China

DOI 10.4108/eai.2-6-2023.2334668

the accuracy of predictions. Liang-Yong Wang[3] proposed crystal characterization based on deep learning image analysis, where mathematical statistics of crystal size and shape are used to predict crystal size expectation and standard deviation. Weijie Wang[4] proposed a continuous crystallization process modelling based on physical information neural network to construct a PINN-based crystallization process to predict the crystal size distribution, and then obtain the optimal constant supersaturation by minimizing the crystal size distribution curve solved by PINN with the expected CSD to calculate the temperature profile of the solution. Sui Sui Chen[5] proposed a study of crystal two-dimensional size growth based on real-time image analysis, using a mask region convolutional neural network and density estimation based on Gaussian kernel function to estimate crystal size. Fuzin Li[6] proposed a study of LSTM-RNN based on LSTM-RN in continuous casting down slag prediction system, using local weighted regression and long short-term memory neural network to predict the down slag moment.

The LSTM, proposed by[7], is a variant of RNN recurrent neural network, which introduces a "memory unit" and can better identify dynamic features through three gate mechanisms The LSTM is a variant of the RNN recurrent neural network, which introduces a "memory unit" that regulates the flow of information through three gate mechanisms, allowing a better identification of its dynamic features[8] . Considering that the crystallization of protection slag is a longer time sequence and is real-time and dynamic, this paper proposes a fusion of convolutional neural network (CNN) and long short-term memory neural network (LSTM) to study the crystallization of protection slag, abandoning the identification of the current moment of protection slag as a feature, but using the current moment and the previous moment or multiple moments of protection slag as features to train the network model. Validation. The experimental results show that the method model can better analyse and predict the dynamic features of crystallization using sequence images, and has higher accuracy and practicality compared with the traditional supervised and unsupervised learning methods.

2 CNN-LSTM BASED CRYSTALLIZATION BEHAVIOR STUDY METHOD

2.1 Data set pre-processing

In the industrial field, the acquisition of industrial data is very limited due to the high cost of machine equipment and sensors, and the difficulty of data collection. Therefore, this paper expands the data set by using data augmentation with the addition of random Gaussian noise on top of the original data. The basic operation is as Eq:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-u)^2}{2\sigma^2}} \quad (1)$$

In order to avoid the CNN-LSTM model consuming a large amount of computational resources when processing high-dimensional input data, this paper uses grayscale processing to reduce the dimensionality of the images. In addition, to improve the performance of the model and speed up the training process, we use the z-score normalization method to preprocess the image data by scaling the range to between 0 and 1, which is calculated as follows:

$$x_{norm} = \frac{x-\mu}{\sigma} \quad (2)$$

2.2 Model Processing Flow

2.2.1 Convolutional Neural Network

CNN, as a typical deep learning model, is more outstanding in the field of image processing, and CNN-based feature extraction converts images into a vector representation in a high-dimensional feature space, which is later used as an input for training prediction, effectively improving the accuracy and efficiency of prediction[9].CNN models can automatically learn the relevant extraction of input data through convolution, pooling and nonlinear operations, and their local connectivity enables the network able to extract local features of images more efficiently, and at the same time, the CNN model greatly reduces the problem of overfitting due to weight sharing. In the convolutional layer, the matrix values in the "field of perception" are scanned by sliding windows and convolutional operations are performed to extract local features.

$$C(i,j) = \sum_{m=1}^{M}\sum_{n=1}^{N} I(i+m, j+n) \cdot K(m,n) \tag{3}$$

Where $C(i,j)$ is the pixel value of the output feature image after convolution, $I(i+m, j+n)$ is the pixel value of the input image, and $K(m,n)$ is the weight value of the convolution kernel.

After obtaining the matrix after convolution, a down sampling operation is performed, and the size of the feature map is reduced, the model parameters are lowered, and the main features of the image are proposed by maximum pooling or mean pooling, etc. in the pooling layer, and the maximum pooling is expressed as

$$P(i,j) = max_{m,n}\big(C(s \cdot i + m, s \cdot j + n)\big) \tag{4}$$

where $P(i,j)$ is the pixel value of the output feature image after pooling, $C(s \cdot i + m, s \cdot j + n)$ is the maximum value within the pooling window, and s is the size of the pooling window.

Finally, by stacking multiple convolutional and pooling layers, higher-level features are extracted and fed to the fully connected layer for multidimensional data downscaling, which is pulled into a one-dimensional vector. This vector can be directly used as the input to the LSTM for model training and prediction.

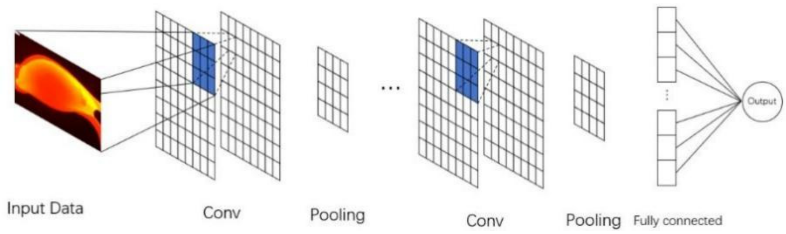

Figure 1: Convolutional Neural Network

2.2.2 Long Short Term Recurrent Memory Network

LSTM is a special kind of recurrent neural network, whose network structure is composed of several small neural networks connected cyclically. Through its cell state and gating mechanism, it can effectively handle long-term dependencies and greatly avoid the problems of gradient disappearance and gradient explosion after the introduction of gating mechanism. Meanwhile, due to the introduction of cell states, the LSTM has the memory ability to capture and store important information in the sequence, which makes the LSTM capable of learning long-term memory. After getting the vector to be predicted, the needed information is first retained from the cell cell by the forgetting gate, which accepts the input from the previous moment and the current time step and outputs a vector between 0 and 1, and then multiplies it element by element with the long-term memory of the previous moment to control the needed information to be retained, as follows:

$$f_{t=}sigmoid\left(U_f h_{t-1} + W_f x_t + b_f\right) \tag{5}$$

where U_f is the forgetting gate weight, h_{t-1} is the hidden layer vector at moment t-1, W_f is the input weight, x_t is the input vector at moment t, and b_f is the bias loop weight.

Next, the information to be memorized is determined by the input gate, and a vector between 0 and 1 is output by the sigmoid function through the memory state of the previous moment and the input of the current moment, with the following equation:

$$i_t = sigmoid(W_i x_t + U_i h_{t-1} + b_i) \tag{6}$$

where W_i is the weight matrix of the input gate, U_i is the input gate weight, and b_i is the bias vector of the input gate.

After the forgetting gate and input gate calculations, the memory state is updated by the vector of their outputs. The memory state is updated using the output of the input gates and the weighting of the candidate memory vectors. The candidate memory state is the multiplication of the outputs of the forgetting gate and the input gate and uses a tanh activation function with the following equation:

$$c_t = f_t c_{t-1} + i_t tan\,h(b_c + W_c x_t + U_c h_{t-1}) \tag{7}$$

ere f_t is the output vector of the forgetting gate, c_{t-1} is the memory state of the previous moment, i_t is the output vector of the input gate, b_c is the bias term of the candidate memory vector, W_c is the weight matrix of the input, x_t is the input vector of the current moment, and U_c is the weight matrix of the hidden state of the previous moment, and the vector is mapped to [-1,1] by the tanh activation function to maintain the range of the memory vector. [10]

Finally, the information output is controlled by the output gate with the following equation:

$$o_t = sigmoid(b_o + W_o x_t + U_o h_{t-1}) \tag{8}$$

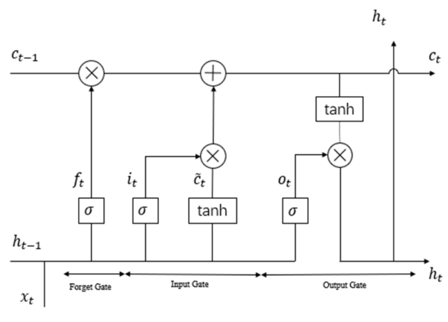

Figure 2: Long Short Term Recurrent Memory Network

2.2.3 Combined CNN-LSTM model

In this paper, a combined CNN-LSTM model is used to study the dynamic features of conservation slag crystals. The CNN and LSTM are connected through a fully connected layer to achieve the prediction of sequence information. After the CNN extracts the static features of the data, the LSTM is then used to model and predict the feature sequences in time, which can effectively deal with the spatio-temporal dependence. The model has both powerful static feature extraction ability and good capability of sequence information processing, and is more suitable for dealing with dynamic feature analysis of crystallization. Its structure diagram is as shown in Figure 3.

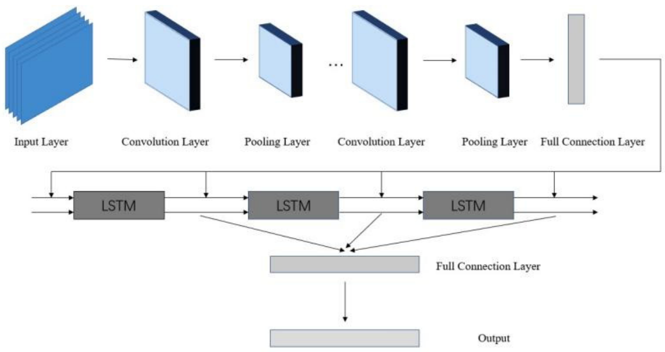

Figure 3: CNN-LSTM Model

2.3 Evaluation indexes

To evaluate the prediction effect of the model, this paper uses the root mean square error (RMSE) and the mean absolute error (MAE)[11] as the evaluation index of the model prediction results, and the smaller these two errors indicate the higher accuracy of the model. The formula is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{9}$$

$$MAE = \frac{\sum_{i=1}^{n}|x_i - y_i|}{n} \tag{10}$$

3 EXPERIMENT AND ANALYSIS OF RESULTS

3.1 Data set and experimental environment

The experiment is conducted on a server with NVIDIA Quadro P2200 graphics card, CUDA version 11.5, and pytorch version 1.9.0. The data set is a total of 562 images from 110s to 671s of protected dregs sequences, and after noise processing, the data set increases to 1124 images. Because the data set has sequential nature, this paper takes out one image every two as the training set, and the remaining part as the test set, of which the test set accounts for about 70% and the test set accounts for about 30%, in line with the conventional data division.

3.2 Model Parameter Setting

Setting the model parameters is not only to define and adjust the structure of the neural network, but also its ability to improve the learning ability, generalization ability and performance of the network according to the task. The deeper the network depth, the better the learning ability of the model, but excessive depth can also lead to overfitting of the model. This model uses the typical ResNet model in CNN, which has a residual connectivity mechanism that allows the network to learn residual information, i.e., the difference between input and output, enabling it to better capture fine-grained features of an image. In this model, we use the model parameters as shown in table 1.

Table 1: Model Parameters

layer	number of layers	activation function
Conv2D	6	RELU
BatchNormalization	5	RELU
Activation	5	RELU
Add	3	RELU

In this study, we get only one feature and a small amount of data, so we do not need a more complex network, and use one LSTM layer for feature learning and one Dense layer for outputting prediction results. The model parameters are as shown in table 2.

Table 2: Model Parameters

Layer	Output Shape	Param
InputLayer	(20,1)	0
LSTM	8	320
Dense	1	9

3.3 Prediction Results and Analysis

In order to better judge the prediction results of the model, this paper compares with CNN-LSTM model by adding SARIMA and ARIMA. With the consistent training and prediction sets, the prediction results are plotted as shown in Figure 4.

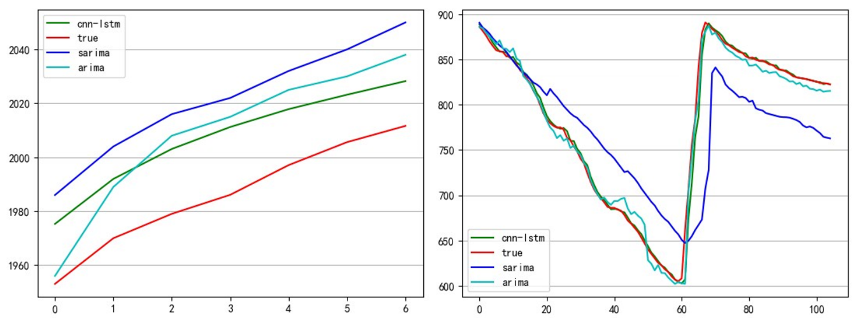

Figure 4: Comparison Results Of SARIMA, ARIMA and CNN-LSTM Models

As seen in Figure 4, the SARIMA model has a weak learning ability for nonlinear data, and the predicted curves of the model deviate more from the true curves, while the predicted curves of the ARIMA model and the CNN-LSTM model deviate less from the true curves, but the volatility of the ARIMA model is larger compared with the CNN-LSTM model. Therefore, the crystallized predicted value curves of the CNN-LSTM model are more consistent with the true value curves than those of the SARIMA and ARIMA models.

As shown in Figure 5, reflects the loss and val_loss of the CNN-LSTM model. loss can be used to measure the error between the predicted and actual results of the model. val_loss changes can determine whether the model is over-fitted or under-fitted. The values of loss and val_loss decrease gradually with the increase of iterations and finally converge to 0. This indicates that the model has a good generalization performance and also shows that the prediction accuracy of the model is also high.

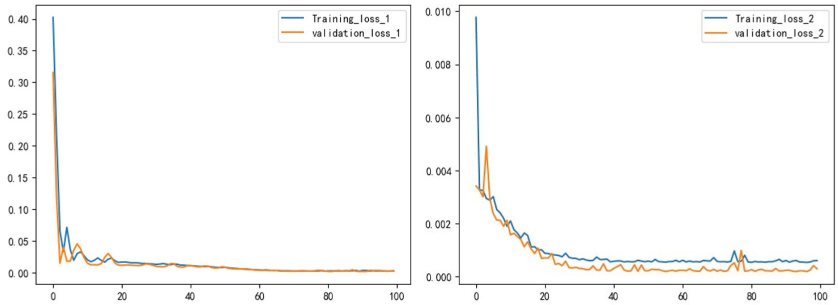

Figure 5.CNN-LSTM Model Loss Results Graph

As shown in table 3 and table 4. It can be seen that the CNN-LSTM model is the lowest for both predictions both MSE and RMSE. So the CNN-LSTM model has a better fit and less error.

Table 3: Melting

Layer	MSE	RMSE
SARIMA	0.00565	0.00697
ARIMA	0.00452	0.00669
CNN-LSTM	0.00337	0.00580

Table 4: Crystallization

Layer	MSE	RMSE
SARIMA	0.01643	0.04054
ARIMA	0.00229	0.05603
CNN-LSTM	0.00125	0.03547

4 SUMMARY

In this paper, a CNN-LSTM based fusion model is proposed for the dynamic features of protection slag crystallization to predict the dynamic features of protection slag crystallization, which can analyse the protection slag crystallization in advance to some extent. The combination of static features extracted by CNN and time-domain features extracted by LSTM, i.e., dynamic features, is carried out to achieve the prediction of dynamic features of protective slag crystallization. And by using RMSE as the evaluation index, and comparing and analysing with SARIMA and ARIMA models, it finally shows that the CNN-LSTM model used in this paper exhibits in addition to better fitting and prediction effects, which provides certain theoretical and technical support for improving steel production efficiency and product quality. It is worth noting that although this paper has achieved an increase in the data set through data enhancement, its sample size is still small, which may affect the generalization and generalization ability of the results, so in further research, we can consider expanding the sample size and increasing the number of network layers and models to a certain extent, so that the model can better handle the analysis of the dynamic characteristics of protective slag crystallization. In future research, increasing the sample size and at the same time increasing the network depth to a certain extent can be considered to further improve the generalization of the model.

REFERENCES

[1] Liu Shaohan. Study on metallurgical behavior of non-reactive continuous casting protection slag under electromagnetic field[D]. North China University of Technology, 2021. DOI:10.27108/d.cnki.ghelu.2021.000957.

[2] Fang HQ, Yan HQ, Chen B et al. Mobile trajectory anomaly detection based on self-coding network[J]. Geoinformation World,2019,26(05):41-44+52.

[3] Wang L.Y.,Zhu Y.L.,Gan C.Y.. Feedback control of crystal properties based on deep learning image analysis[J]. Journal of Northeastern University (Natural Science Edition),2022,43(12):1688-1693.

[4] Wang Weijie. Modeling and optimization of continuous crystallization process based on physical information neural network[D]. Dalian University of Technology, 2021. DOI:10.26991/d.cnki.gdllu.2021.000281.

[5] Chen Sui Su. Research on two-dimensional dimensional growth of crystals based on real-time image analysis[D]. Dalian University of Technology,2019.DOI:10.26991/d.cnki.gdllu.2019.001506.

[6] Li Fujin,Liu Shangyu,Shi Tao.Application of LSTM-RNN in continuous casting down slag prediction system[J]. Machine Design and Manufacture,2022,No.371(01):181-183+188.DOI:10.19356/j.cnki.1001-3997.2022.01.006.

[7] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. neural Computation, 9(8), 1735-1780. doi: 10.1162/neco.1997.9.8.1735

[8] Swaymprabha Alias Megha Mane, Arundhati Shinde,StressNet: Hybrid model of LSTM and CNN for stress detection from electroencephalogram signal (EEG),Results in Control and Optimization,2023,100231,ISSN 2666-7207,https://doi.org/10.1016/j.rico.2023.100231.

[9] Bi D.J.,Wei S.L.,Zhao T. et al. Sentiment analysis model of online reviews based on convolutional neural network[J]. Journal of Hebei Science and Technology Teachers College,2019,33(02):41-47.

[10] Tao D,Yao Y,Wu Zunxi et al. A recommendation model of interest points with user preferences and time series[J]. Small Microcomputer Systems,2022,43(03):582-588.DOI:10.20009/j.cnki.21-1106/TP.2021-0088.

[11] Jiaxuan Zhang, Shunyong Li,Air quality index forecast in Beijing based on CNN-LSTM multi-model,Chemosphere,Volume 308, Part 1,2022,136180. ISSN 0045-6535, https://doi.org/10.1016/j.chemosphere.2022.136180.