Supplement for

TrCSL: A Transferred CNN-SE-LSTM Model for High-accuracy Quantitative Analysis of Laser-induced Breakdown Spectroscopy with Small Samples

Shengjie Ma[a,b,c], Shilong Xu[a,b,c], Congyuan Pan[d], Jiajie Fang[a,b,c], Fei Han[a,b,c], Yuhao Xia[a,b,c], Wanying Ding[a,b,c], Youlong Chen[a,b,c] and Yihua Hu*[a,b,c]*

a. State Key Laboratory of Pulsed Power Laser Technology, National University of Defense Technology, Hefei 230037, People's Republic of China

b. Key Laboratory of Electronic Restriction of Anhui Province, National University of Defense Technology, Hefei 230037, People's Republic of China

c. Advanced Laser Technology Laboratory of Anhui Province, National University of Defense Technology, Hefei 230037, People's Republic of China

d. Hefei GStar Intelligent Control Technical Co.,Ltd, Hefei, 230037, People's Republic of China.

Corresponding Author

*Email: skl_hyh@163.com; xushi1988@nudt.edu.cn

1. Process of samples preparation.

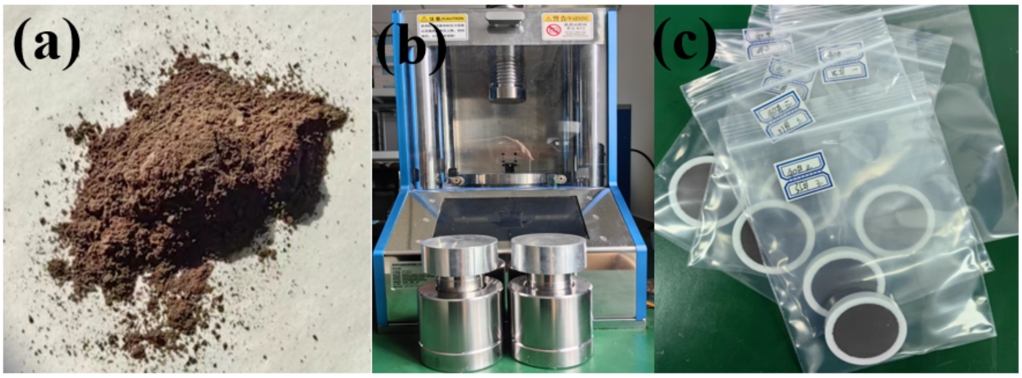

Fig.S1 (a) is the powdery steel slag sample, (b) is the tablet press produced by Shanghai Xinnuo Instrument Equipment Co., Ltd., and (c) is sample after tablet pressing.

2. Some information of the samples

Tab.S1 The concentration of the major components in steel slag samples (the source data of TrCSL model).

No.sample	TFe	SiO_2	CaO	MgO	Al_2O_3	S	P_2O_5	TiO_2	MnO
1	18.12	11.55	43.15	8.53	1.97	0.08	2.47	1.12	3.56
2	16.78	13.49	42.26	8.02	1.71	0.06	3.07	1.16	3.91
3	18.43	11.53	41.7	8.83	2	0.08	2.55	1.2	3.77
4	25.97	10.21	37.41	7.77	1.91	0.09	2.33	1.13	3.54
5	17.08	13.72	42.24	7.93	1.7	0.07	3.11	1.19	3.98
6	19.49	11.42	41.05	8.66	2.6	0.07	2.49	1.17	3.8
7	18.55	13.38	41.33	7.36	1.7	0.06	3.02	1.2	4.07
8	17.65	13.24	42.06	8.16	1.74	0.06	3.05	1.13	3.98
9	20.52	11.78	40.93	8.08	1.85	0.06	2.41	1.05	3.66
10	17.88	13.89	41.96	7.84	1.83	0.07	3.07	1.26	4.16
11	17.86	13.49	42.01	7.82	1.85	0.05	3.03	1.18	4.09
12	19.12	12.09	41.22	8.65	2.14	0.08	2.55	1.2	3.71
13	20.45	10.67	41.35	7.67	1.93	0.08	2.46	1.12	3.68
14	32.44	9.22	31.51	2.82	1.96	0.11	1.66	0.73	3.89
15	21.51	10.77	39.71	7.89	1.69	0.06	2.41	1.05	3.97
16	29.29	10.71	33.54	3.18	2.5	0.1	1.97	0.89	4.35
17	31.54	10.05	31.29	2.79	2.62	0.1	1.76	0.9	4.07
18	23.04	12.05	38.13	6.65	1.73	0.05	2.64	1.12	3.68
19	18.23	13.22	41.26	7.82	1.92	0.05	3.05	1.16	3.91
20	18.39	13	41.24	7.83	1.93	0.06	2.96	1.16	3.86
21	17.43	13.04	42.21	7.55	1.95	0.07	3.01	1.18	3.87
22	22.27	10.86	39.69	8.34	1.78	0.05	2.48	1.16	3.85
23	21.29	10.65	40.99	8.31	1.71	0.06	2.52	1.14	3.81
24	17.73	13.71	42.87	7.49	1.8	0.06	3.16	1.22	4.07
25	22.68	10.03	41.12	8.19	1.49	0.08	2.63	0.93	3.78
26	19.87	12.5	41.69	7.47	1.74	0.06	3.12	1.17	3.97
27	22.26	12.08	39.11	6.83	1.73	0.06	2.82	1.23	3.87
28	31.99	9.54	30.87	2.73	2.34	0.11	1.76	0.9	4.21
29	20.2	10.66	42.5	8.38	1.53	0.07	2.71	0.99	3.91
30	30.49	10.01	32.41	2.81	2.5	0.1	1.9	0.93	4.42
31	29.47	10.27	33	2.81	2.65	0.1	1.94	0.96	4.42
32	35.56	9.46	28.88	2.79	2.41	0.09	1.67	0.88	3.89
33	30.93	9.64	31.88	2.68	2.46	0.1	1.75	0.91	4.11
34	29.51	10.42	32.76	3.16	2.6	0.11	1.86	0.89	4.29
35	29.31	10.38	32.41	3.02	2.5	0.1	1.87	0.91	4.38
36	20.98	10.83	40.31	8.16	1.53	0.07	2.47	1.04	3.86
37	21.58	12.21	38.77	7.64	1.64	0.06	2.74	1.17	3.71
38	29.47	10.21	33.06	3.09	2.39	0.1	1.91	0.9	4.29
39	29.72	10.3	33.03	3.01	2.51	0.11	1.87	0.9	4.28

40	31.81	9.95	31.58	2.81	2.46	0.1	1.81	0.88	4.2
41	17.45	12.52	42.70	8.27	1.84	0.07	2.77	1.14	3.73
42	18.27	11.54	42.42	8.68	1.98	0.08	2.51	1.16	3.66
43	22.04	10.88	40.28	8.15	1.94	0.08	2.4	1.12	3.55
44	17.6	12.63	42.69	8.23	1.83	0.07	2.79	1.15	3.77
45	18.80	11.48	42.1	8.59	2.28	0.07	2.48	1.14	3.68
46	18.33	12.46	42.24	7.94	1.83	0.07	2.7	1.16	3.81
47	17.88	12.39	42.60	8.34	1.85	0.07	2.76	1.12	3.77
48	19.32	11.66	42.04	8.30	1.91	0.07	2.44	1.08	3.61
49	18	12.72	42.55	8.18	1.9	0.07	2.77	1.19	3.86
50	17.99	12.52	42.58	8.17	1.91	0.06	2.75	1.15	3.82
51	18.62	11.82	42.18	8.59	2.05	0.08	2.51	1.16	3.63
52	19.28	11.11	42.25	8.1	1.95	0.08	2.465	1.12	3.62
53	25.28	10.38	37.33	5.67	1.96	0.07	2.065	0.925	3.72
54	19.81	11.16	41.43	8.21	1.83	0.07	2.44	1.085	3.76
55	23.70	11.13	38.34	5.85	2.23	0.09	2.22	1.005	3.95
56	24.83	10.8	37.22	5.66	2.29	0.09	2.11	1.01	3.81
57	20.58	11.8	40.64	7.59	1.85	0.06	2.55	1.12	3.62
58	18.17	12.38	42.20	8.17	1.94	0.06	2.76	1.14	3.73
59	18.25	12.27	42.19	8.18	1.95	0.07	2.71	1.14	3.71
60	17.77	12.29	42.68	8.04	1.96	0.07	2.74	1.15	3.71
61	20.19	11.20	41.42	8.43	1.87	0.06	2.47	1.14	3.70
62	19.70	11.1	42.07	8.42	1.84	0.07	2.49	1.13	3.68
63	17.92	12.63	43.01	8.01	1.88	0.07	2.81	1.17	3.81
64	20.4	10.79	42.13	8.36	1.73	0.08	2.55	1.02	3.67
65	18.99	12.02	42.42	8	1.85	0.07	2.79	1.14	3.76
66	20.19	11.81	41.13	7.68	1.85	0.07	2.64	1.17	3.71
67	25.05	10.54	37.01	5.63	2.15	0.09	2.11	1.01	3.88
68	19.16	11.10	42.82	8.45	1.75	0.075	2.59	1.05	3.73
69	24.30	10.78	37.78	5.67	2.23	0.09	2.18	1.02	3.99
70	23.79	10.91	38.07	5.67	2.31	0.09	2.20	1.04	3.99
71	26.84	10.50	36.01	5.66	2.19	0.08	2.07	1	3.72
72	24.52	10.59	37.51	5.60	2.21	0.09	2.11	1.01	3.83
73	23.81	10.98	37.95	5.84	2.28	0.09	2.16	1.00	3.92
74	23.71	10.96	37.78	5.77	2.23	0.09	2.17	1.01	3.97
75	19.55	11.19	41.73	8.34	1.75	0.07	2.47	1.08	3.71
76	19.85	11.88	40.96	8.08	1.80	0.07	2.60	1.14	3.63
77	23.79	10.88	38.10	5.81	2.18	0.09	2.19	1.01	3.92
78	23.92	10.92	38.09	5.77	2.24	0.09	2.17	1.01	3.92
79	24.96	10.75	37.36	5.67	2.21	0.09	2.14	1	3.88
80	17.60	12.51	41.98	8.42	1.85	0.07	2.81	1.18	3.84
81	21.37	11.85	39.83	7.89	1.81	0.07	2.7	1.14	3.72
82	16.93	13.60	42.25	7.97	1.70	0.06	3.09	1.17	3.94
83	18.13	12.45	41.65	8.34	2.15	0.06	2.78	1.16	3.85

84	17.66	13.43	41.79	7.69	1.70	0.06	3.04	1.18	3.99
85	17.21	13.36	42.16	8.09	1.72	0.06	3.06	1.14	3.94
86	18.65	12.63	41.59	8.05	1.78	0.06	2.74	1.10	3.78
87	17.33	13.69	42.11	7.93	1.77	0.06	3.07	1.21	4.03
88	17.32	13.49	42.13	7.92	1.78	0.05	3.05	1.17	4
89	17.95	12.79	41.74	8.33	1.92	0.07	2.81	1.18	3.81
90	18.615	12.08	41.80	7.84	1.82	0.07	2.76	1.14	3.79
91	24.61	11.35	36.88	5.42	1.83	0.08	2.36	0.94	3.9
92	19.14	12.13	40.98	7.95	1.7	0.06	2.74	1.10	3.94
93	23.035	12.1	37.9	5.6	2.10	0.08	2.52	1.02	4.13
94	24.16	11.77	36.77	5.40	2.16	0.08	2.41	1.03	3.99
95	19.91	12.77	40.19	7.33	1.72	0.05	2.85	1.14	3.79
96	17.50	13.35	41.76	7.92	1.81	0.05	3.06	1.16	3.91
97	17.58	13.24	41.75	7.92	1.82	0.06	3.01	1.16	3.88
98	17.10	13.26	42.23	7.78	1.83	0.06	3.04	1.17	3.89
99	19.52	12.17	40.97	8.18	1.745	0.05	2.77	1.16	3.88
100	19.03	12.07	41.62	8.16	1.71	0.06	2.79	1.15	3.86

Tab.S2 The concentration of the major components in carbon steel samples (the target data of TrCSL model).

No.sample	Cu	S	Fe	Si	Ni	Mn	Cr	C	P
1	0.0004	0.00024	99.98216	0.0018	0.0012	0.0008	0.0003	0.002	0.0014
2	0.0088	0.0031	99.8253	0.016	0.016	0.082	0.022	0.0064	0.0066
3	0.015	0.025	98.3873	0.24	0.015	0.72	0.021	0.483	0.01
4	0.197	0.027	98.5133	0.202	0.07	0.554	0.12	0.196	0.009
5	0.016	0.0024	98.2761	0.22	0.038	0.22	0.261	0.92	0.015
6	0.081	0.059	98.9	0.193	0.032	0.459	0.083	0.165	0.028
7	0.02	0.011	99.0371	0.241	0.015	0.459	0.024	0.183	0.0099
8	0.012	0.032	98.649	0.286	0.014	0.583	0.122	0.273	0.029
9	0.355	0.031	97.503	0.48	0.12	0.87	0.28	0.31	0.051
10	0.0046	0.00167	99.90373	0.0089	0.0086	0.0414	0.01115	0.0042	0.004
11	0.0077	0.01262	99.18473	0.1209	0.0081	0.3604	0.01065	0.2425	0.0057
12	0.0987	0.01362	99.24773	0.1019	0.0356	0.2774	0.06015	0.099	0.0052
13	0.0082	0.00132	99.12913	0.1109	0.0196	0.1104	0.13065	0.461	0.0082
14	0.0119	0.01405	99.1063	0.128	0.0155	0.401	0.0215	0.2447	0.0083
15	0.1029	0.01505	99.1693	0.109	0.043	0.318	0.071	0.1012	0.0078
16	0.0124	0.00275	99.0507	0.1180	0.027	0.151	0.1415	0.4632	0.0108
17	0.106	0.026	98.4503	0.221	0.0425	0.637	0.0705	0.3395	0.0095
18	0.0155	0.0137	98.3317	0.23	0.0265	0.47	0.141	0.7015	0.0125
19	0.1065	0.0147	98.3947	0.211	0.054	0.387	0.1905	0.558	0.012
20	0.0505	0.035	98.96855	0.217	0.0235	0.459	0.0535	0.174	0.01895

Table S3 Spectral lines of the carbon steel samples.

Element	Spectral line
C	344.471 nm, 414.626 nm, 426.902 nm, 476.253 nm, 489.065 nm, 526.894 nm, 538.034 nm
Cu	356.612 nm, 361.420 nm, 361.420 nm, 370.052 nm, 372.076 nm, 386.046 nm, 406.263 nm, 407.557 nm, 424.8956 nm, 427.510 nm, 441.555 nm, 452.512 nm

Mn	344.198 nm, 357.788 nm, 361.926 nm, 379.02 nm, 383.386 nm, 384.108 nm, 404.136 nm, 407.028 nm, 423.514 nm, 425.766 nm, 441.488 nm, 446.468 nm, 447.279 nm
Cr	357.869 nm, 363.146 nm, 390.292 nm, 392.102 nm, 392.864 nm, 399.112 nm, 425.435 nm, 427.48 nm, 428.972 nm, 438.498 nm, 449.686 nm, 452.647 nm, 526.572 nm, 532.834 nm

3. Structure of LSTM

The LSTM model is mainly composed of the memory cell, forget gate, input gate, and output gate.

The purpose of the forget gate is to decide what information to discard. The output f_t of the forgetting gate can be expressed as

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right), \tag{S1}$$

The forget gate obtains the output with the range of [0,1] and decides how much information to retain according to the previous hidden state h_{t-1} and the current input information x_t. 1 means all the information remained and 0 means filter all the information out.

The input gate's role is to update the state of the memory cell. The input gate consists of a sigmoid and tanh activation function, where the sigmoid function determines which information to discard, and the tanh function generates new candidate information.

$$i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right), \tag{S2}$$

$$\tilde{C}_t = \tanh\left(W_C\left[h_{t-1}, x_t\right] + b_C\right), \tag{S3}$$

And the memory cell status changes to

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \tag{S4}$$

The output gate also consists of a sigmoid and a tanh activation function. The sigmoid function determines what information should be retained, while the tanh function handles the state of the memory unit in preparation for output.

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right), \tag{S5}$$

$$h_t = o_t * \tanh\left(C_t\right), \tag{S6}$$

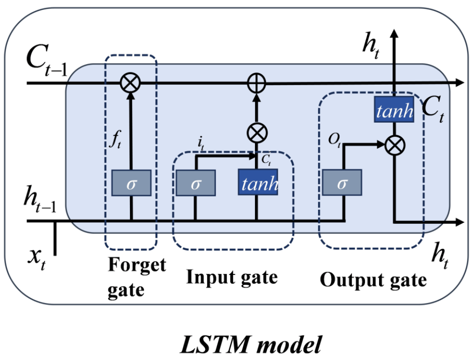

Fig.S2 The structure of LSTM model.

3. Hyperparameters of the model

Kernel size and number. Each convolutional kernel is responsible for extracting local features from the LIBS spectrum. A multitude of kernels work in concert to comprehensively extract the main features from the input samples. The size and number of the convolutional kernels collectively determine the breadth and diversity of features that the network can extract. Larger kernels offer a greater receptive field, thereby more effectively capturing global features. However, this also implies a higher demand for computational resources, which may negatively impact computational performance.

Stride. The stride is a pivotal parameter in convolutional computations, defining the number of pixels the convolutional kernel moves across the input data with each slide (as shown in Fig.5(a), the stride is set to 2). The magnitude of the stride directly affects the granularity of the convolutional operation. Taking the illustration in Fig.5 as an example, for a 2×2 kernel, if the stride is set to 1, the neighboring receptive fields will partially overlap, which is beneficial for capturing continuous feature variations. Conversely, when the stride is set to 2, there is no overlap between the receptive fields, allowing for faster coverage of the entire input but potentially overlooking some local details. If the stride is increased to 3, a gap of one pixel width will occur between the neighboring receptive fields, which may result in the omission of some important original information during the feature extraction process.

Padding. Padding refers to the process of adding extra rows or columns, typically filled with zeros, around the input data. This allows for the artificial adjustment of the output shape based on the given input and kernel size, facilitating the prediction and alteration of the output dimensions as desired. This technique is highly useful when designing and constructing neural network architectures, as it enables precise planning of the network's structure by allowing us to forecast and control the output shape of each layer.

Activation function. One of the roles of activation function is to enable the model to effectively handle nonlinear problems. The convolutional operation shown in Fig.5(a) and pooling operation in Fig.5(b) are essentially linear transformations. No matter how many layers the neural network has, if it contains only linear operations, the final output of the network can be represented as a linear combination of inputs, meaning that the network is essentially still a linear model. To enhance the ability to deal with complex nonlinear tasks, the introduction of activation functions is crucial, which helps the model to better adapt to diverse data and improve its generalization capability. Common activation functions include sigmoid, tanh, ReLU, and so on.

Batch size. Batch size refers to the number of samples used in each parameter update of the neural network model. It significantly determines the model's training performance and speed. Generally, a larger batch size is more efficient for training data, allowing the model to converge more rapidly. However, this may also restrict the model's generalization ability, as the model can easily become trapped in sharp local minima and struggle to escape. On the other hand, a smaller batch size facilitates the model's exploration of a broader parameter space, potentially finding better global minimum points. But this also increases computational load and training time, demanding higher specifications from hardware resources.

Learning rate. The learning rate defines the step size used for parameter adjustments. The magnitude of the learning rate directly affects the model's convergence speed and the effectiveness of the training process. A well-chosen learning rate can lead to more stable training and a faster approach to the minimum of the loss function. However, setting the learning rate too high, while seemingly accelerating convergence with larger parameter updates, can actually result in an unstable training process or failure to converge effectively. Therefore, selecting an appropriate learning rate is essential for ensuring the efficiency and stability of model training.

Epoch. Epoch refers to the process in model training where all training samples go through one complete forward propagation followed by a backward propagation.

4. Influence analysis of hyperparameters on model performance

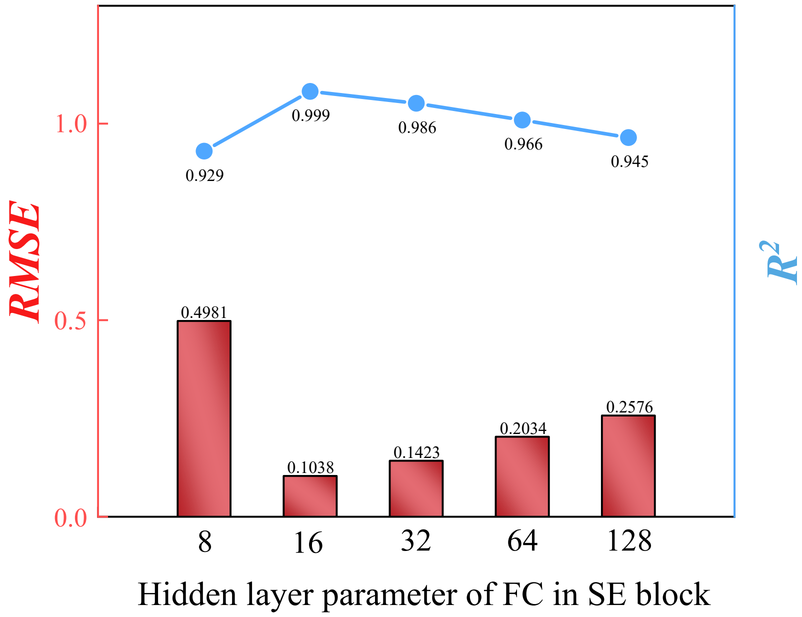

Fig.S3 Performance of 2D CNN-SE-LSTM with different Hidden layer parameter of FC in SE block.

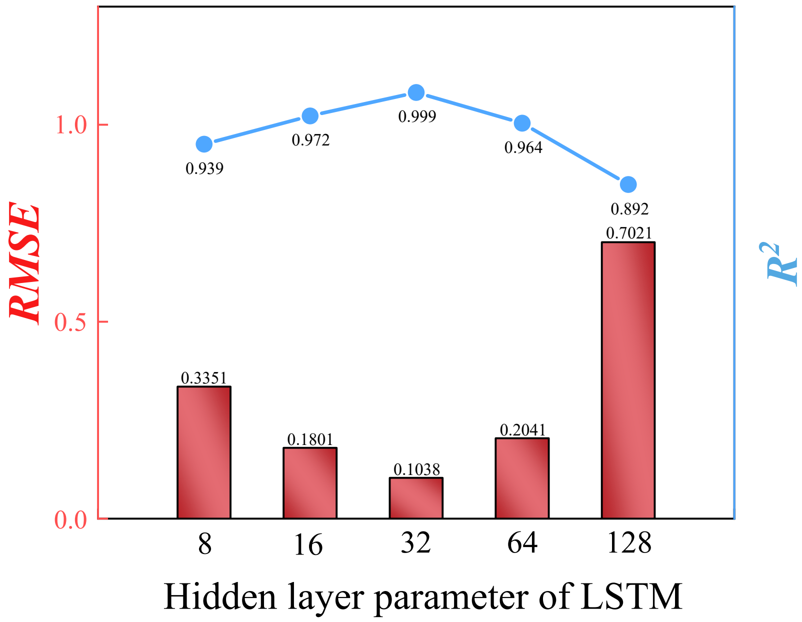

Fig.S4 Performance of 2D CNN-SE-LSTM with different hidden layer parameter of LSTM.

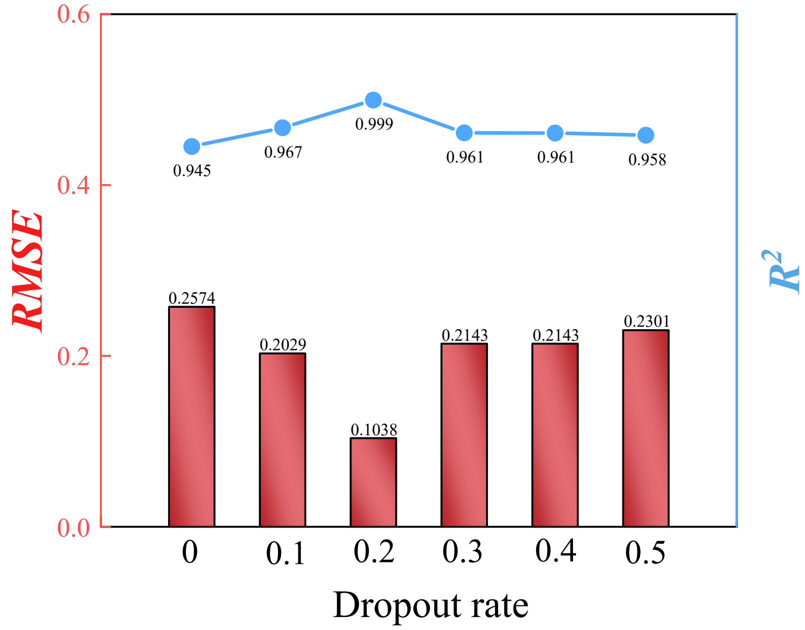

Fig.S5 Performance of 2D CNN-SE-LSTM with different dropout rate.

Fig.S5 compares the results as the Dropout rate varies from 0 to 0.5. The experimental results indicate that when the Dropout rate is set to 0.2, the prediction performance is optimal, and compared to the non-regularized model (Dropout rate=0), the prediction results showed a reduction of 0.1536 in RMSE and an increase of 0.0543 in R^2.

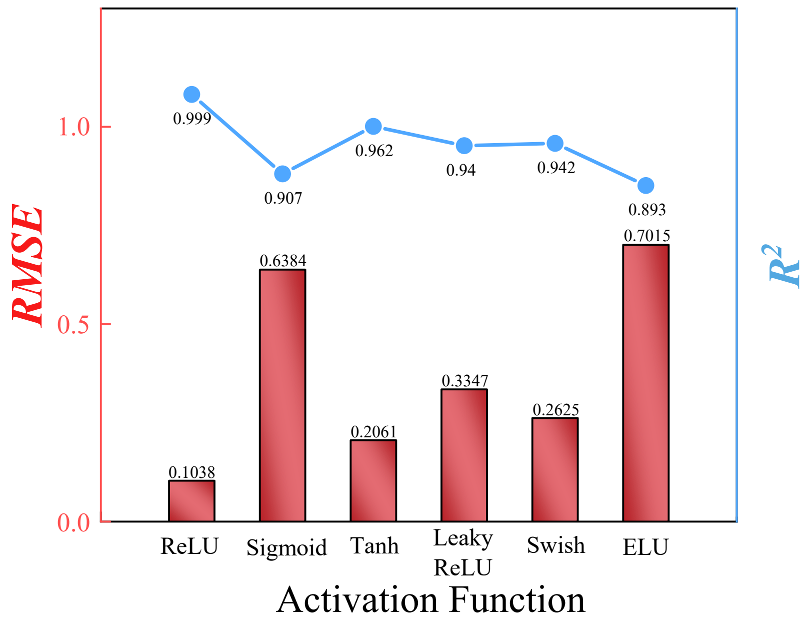

Fig.S6 Performance of 2D CNN-SE-LSTM with different activation functions.

According to the results shown in Fig.S6, the ReLU function demonstrated the most outstanding performance in prediction. Compared to the Sigmoid and Tanh function, ReLU can maintain gradient non-decay within the positive range due to its non-saturating nature, effectively alleviating the problem of vanishing gradients. At the same time, the ReLU function avoids the introduction of additional parameters. Furthermore, we assessed the impact of some improved ReLU functions on the results such as the Leaky ReLU, Swish, and ELU function. Regrettably, these three functions did not accept the expected ideal results.

5. Spectral fluctuations analysis

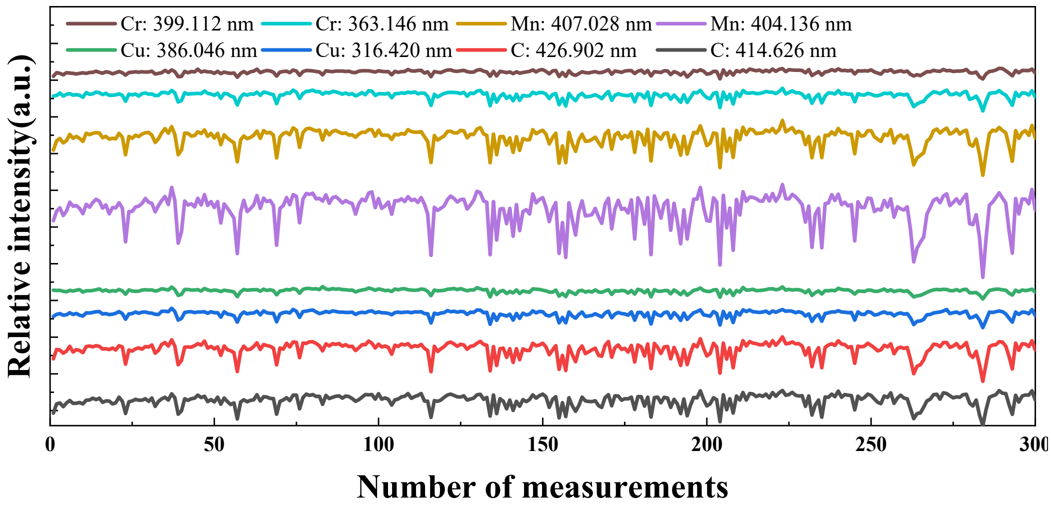

Fig.S7 The intensity fluctuation degree of Cr: 399.12 nm, Cr: 363.146 nm, Cu: 386.046 nm, Cu: 316.420 nm, Mn: 407.028 nm, Mn: 404.136 nm, C: 426.902 nm, and C: 414.626 nm for Sample No.1.